IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLORADO

Civil Action No. 11-cv-01389-WJM-KLM

GENETIC TECHNOLOGIES LIMITED,
an Australian corporation,

Plaintiff,

v.

AGILENT TECHNOLOGIES, INC., a Delaware corporation;
BRISTOL-MYERS SQUIBB COMPANY, a Delaware corporation;
EUROFINS STA LABORATORIES, INC., a Colorado corporation;
GENESEEK, INC., a Nebraska corporation;
GLAXOSMITHKLINE LLC, a Delaware corporation;
HOLOGIC, INC., a Delaware corporation;
MERIAL L.L.C., a Delaware limited liability company;
PFIZER INC., a Delaware corporation; and
454 LIFE SCIENCES CORPORATION, a Delaware corporation;

Defendants.

FIRST AMENDED COMPLAINT WITH JURY DEMAND

Plaintiff Genetic Technologies Limited ("GTG") files this Complaint against Defendants Agilent Technologies, Inc. ("Agilent"), Bristol-Myers Squibb Company ("Bristol-Myers"), Eurofins STA Laboratories, Inc. ("ESTA"), GeneSeek, Inc. ("GeneSeek"), GlaxoSmithKline LLC ("GSK"), Hologic, Inc. ("Hologic"), Merial L.L.C. ("Merial"), Pfizer Inc. ("Pfizer"), and 454 Life Sciences Corporation ("454") (hereinafter referred to collectively as "Defendants" unless otherwise specified) alleging as follows:

I. THE PARTIES

1. Plaintiff GTG is an Australian corporation with a principal place of business in Victoria, Australia.

2. Upon information and belief, Agilent is a corporation organized and existing under the laws of the state of Delaware, with its principal place of business located at 5301 Stevens Creek Boulevard, Santa Clara, California 95051. Agilent can be served with process through its registered agent, The Corporation Trust Company, Corporation Trust Center, 1209 Orange Street, Wilmington, Delaware 19801.

3. Upon information and belief, Bristol-Myers is a corporation organized and existing under the laws of the state of Delaware, with its principal place of business located at 345 Park Avenue, New York, New York 10154. Bristol-Myers can be served with process through its registered agent, The Corporation Trust Company, Corporation Trust Center, 1209 Orange Street, Wilmington, Delaware 19801.

4. Upon information and belief, ESTA is a corporation organized and existing under the laws of the state of Colorado, with its principal place of business located at 1821 Vista View Drive, Longmont, Colorado 80504. ESTA can be served with process through its registered agent, National Corporate Research, Ltd., 12649 West Warren Avenue, Lakewood, Colorado 80228.

5. Upon information and belief, GeneSeek is a corporation organized and existing under the laws of the state of Nebraska, with its principal place of business located at 4665 Innovation Drive, Suite 120, Lincoln, Nebraska 68521. GeneSeek can be served with process at its principal place of business.

6. Upon information and belief, GSK is a company organized and existing under the laws of the state of Delaware, with its principal place of business located at One Franklin Place, 200 North 16th Street, Philadelphia, Pennsylvania 19106. GSK can be served with process through its registered agent, Corporation Service Company, 2711 Centerville Road, Suite 400, Wilmington, Delaware 19808.

7. Upon information and belief, Hologic is a corporation organized and existing under the laws of the state of Delaware, with its principal place of business located at 35 Crosby Drive, Bedford, Massachusetts 01730. Hologic can be served with process through its registered agent, Corporation Service Company, 2711 Centerville Road, Suite 400, Wilmington, Delaware 19808.

8. Upon information and belief, Merial is a limited liability organized and existing under the laws of the state of Delaware, with its principal place of business located at 3239 Satellite Boulevard, Duluth, Georgia 30096. Merial is a joint venture between Merck & Co. and Sanofi-Aventis U.S. Merial can be served with process through its registered agent, The Corporation Trust Company, Corporation Trust Center, 1209 Orange Street Wilmington, Delaware 19801.

9. Upon information and belief, Pfizer is a corporation organized and existing under the laws of the state of Delaware, with its principal place of business located at 235 East 42nd Street, New York, New York 10017-5755. Pfizer can be served with process through its registered agent, The Corporation Trust Company, Corporation Trust Center, 1209 Orange Street, Wilmington, Delaware 19801.

10. 454 is organized and existing under the laws of the state of Delaware with its principal place of business located at 1 Commercial Street, Branford, Connecticut, 06405. 454 can be served with process through its registered agent The Corporation Trust Company, Corporation Trust Center 1209 Orange Street, Wilmington, Delaware 19801.

II. JURISDICTION AND VENUE

11. This Court has exclusive jurisdiction of this action for patent infringement pursuant to 28 U.S.C. § 1338(a).

12. This Court has jurisdiction over the subject matter of this action pursuant to 28 U.S.C. §§ 1331 and 1338(a).

13. Venue is proper in this judicial district pursuant to 28 U.S.C. §§ 1391 and 1400.

14. Upon information and belief, Defendants each have minimum contacts with this judicial district such that this forum is a fair and reasonable one. Defendants have also each committed such purposeful acts and/or transactions in Colorado that they reasonably knew and/or expected that they could be hauled into court as a future consequence of such activity. Upon information and belief, Defendants have also transacted and/or, at the time of the filing of this Amended Complaint, are transacting business within the District of Colorado. For these reasons, personal jurisdiction exists over all Defendants and venue over this action is proper in this Court under 28 U.S.C. §§ 1391(b) and (c) and 28 U.S.C. § 1400(b).

III. THE PATENT-IN-SUIT

15. On March 18, 1997, United States Patent No. 5,612,179 ("the '179 Patent") was duly and legally issued for an "Intron Sequence Analysis Method for Detection of Adjacent

Locus Alleles as Haplotypes." A true and correct copy of the '179 Patent is attached as Exhibit A.

16. GTG is the owner of the '179 Patent by assignment from Genetype AG, who was originally assigned the technology by the inventor Dr. Malcolm Simons, with the exclusive right to enforce and collect damages for infringement of the '179 Patent during all relevant time periods.

17. The '179 Patent generally relates to methods of analysis of non-coding DNA sequences.

18. The Abstract of the '179 Patent relevantly provides:

> The present invention provides a method for detection of at least one allele of a genetic locus and can be used to provide direct determination of the haplotype. The method comprises amplifying genomic DNA with a primer pair that spans an intron sequence and defines a DNA sequence in genetic linkage with an allele to be detected. The primer-defined DNA sequence contains a sufficient number of intron sequence nucleotides to characterize the allele. Genomic DNA is amplified to produce an amplified DNA sequence characteristic of the allele. The amplified DNA sequence is analyzed to detect the presence of a genetic variation in the amplified DNA sequence such as a change in the length of the sequence, gain or loss of a restriction site or substitution of a nucleotide. The variation is characteristic of the allele to be detected and can be used to detect remote alleles.

19. Independent Claims 1 and 26 of the '179 Patent read:

> 1. A method for detection of at least one coding region allele of a multi-allelic genetic locus comprising: a) amplifying genomic DNA with a primer pair that spans a non-coding region sequence, said primer pair defining a DNA sequence which is in genetic linkage with said genetic locus and contains a sufficient number of non-coding region sequence nucleotides to produce an amplified DNA sequence characteristic of said allele; and b) analyzing the amplified DNA sequence to detect the allele.
>
> 26. A DNA analysis method for determining coding region alleles of a multi-allelic genetic locus comprising identifying sequence polymorphisms characteristic of the alleles, wherein said sequence polymorphisms characteristic

of the alleles are present in a non-coding region sequence, said non-coding region sequence being not more than about two kilobases in length.

20. The '179 Patent is presumed valid and enforceable pursuant to 35 U.S.C. § 282.

21. The '179 Patent was previously asserted by GTG in the matter of *Genetic Technologies Ltd. v. Applera Corp.*, Case No. C 03-1316-PJH, in the United States District for the Northern District of California (the "Applera Action"). The Applera Action was ultimately settled with Applera Corporation taking a license to the '179 Patent, among others.

22. The '179 Patent was the subject of a declaratory judgment action initiated by Monsanto in the matter of *Monsanto Company v. Genetic Technologies Ltd.*, Case No. 06-cv-00989-HEA, in the United States District Court for the Eastern District of Missouri, Eastern Division (the "Monsanto Action"). That Monsanto Action was ultimately settled. Monsanto has now taken three licenses to the '179 Patent, among others.

23. The '179 Patent was most recently asserted by GTG in the matter of *Genetic Technologies Ltd. v. Beckman Coulter, Inc., et al*, Case No. 10-cv-0069-BBC, in the United States District Court for the Western District of Wisconsin (the "Beckman Coulter Action"). The Beckman Coulter Action was resolved with at least Beckman Coulter, Inc., Gen-Probe, Inc., Interleukin Genetics Incorporated, Molecular Pathology Laboratory Network, Inc., Orchid Cellmark, Inc., Pioneer Hi-Bred International, Inc., and Sunrise Medical Laboratories, Inc. all taking a license to the '179 Patent, among others. GTG has secured over $14.5 million in licensing revenue since the filing of the Beckman Coulter Action in 2010.

24. In addition to the licenses identified in the preceding paragraphs, the '179 Patent and related patents have been licensed to at least the following entities: AgResearch Ltd.; ARUP Laboratories, Inc.; Australian Genome Research Facility Ltd.; Bio Reference Laboratories

(subsidiary GeneDx); Bionomics Ltd.; BioSearch Technologies Inc.; Pfizer Animal Health; C Y O'Connor ERADE Village Foundation (incorporating the Immunogenetics Research Foundation and the Institute of Molecular Genetics and Immunology Incorporated); Crop and Food Research Ltd.; DNA Diagnostics Ltd.; General Electric Co. and its subsidiary GE Healthcare Bio-Sciences Corp.; Genosense Diagnostics GmbH; Genzyme Corp.; Innogenetics N.V.; Kimball Genetics, Inc.; Laboratory Corporation of America Holdings, Inc.; Livestock Improvement Corporation Ltd.; MetaMorphix, Inc.; Millennium Pharmaceuticals Inc.; Myriad Genetics, Inc.; Nanogen, Inc.; New Zealand Blood Service; Optigen, L.L.C.; Ovita Ltd.; Perlegen Sciences, Inc.; Prometheus Laboratories Inc.; Qiagen, Inc.; Quest Diagnostics Inc.; Sciona, Inc.; Sequenom, Inc.; Syngenta Crop Protection AG; Thermo Fisher Scientific Inc.; TIB MOLBIOL Syntheselabor GmbH; Tm Bioscience Corporation; Gen-Probe, Inc.; and others.

25. Certain claims of the '179 Patent, including Claim 26, were recently subjected to an ex parte reexamination before the United States Patent and Trademark Office ("USPTO") that was initiated by an unknown entity. On February 4, 2010, the USPTO issued a Notice of Intent to Issue Ex Parte Reexamination Certificate indicating that the subject claims were confirmed as valid without amendment. A true and correct copy of that Reexamination Certificate is attached as Exhibit B.

26. The '179 Patent expired on March 9, 2010. However, GTG remains entitled to collect damages for past infringement occurring during the term of the '179 Patent pursuant to 35 U.S.C. §§ 284 and 286.

IV. DEFENDANTS' INFRINGEMENT

27. Upon information and belief, and as further described below, Defendants have manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products and/or services in the United States that infringed one or more claims of the '179 Patent; and/or Defendants have induced and/or contributed to the infringement of one or more of the claims of the '179 Patent by others in the United States.

A. AGILENT

28. Agilent acquired Stratagene Corporation ("Stratagene") in June of 2007, and Stratagene became part of Agilent's Bio-Analytical Measurements business group. According to Stratagene marketing materials, "[i]t offers products in various categories, including amplification, cloning, nucleic acid analysis, quantitative PCR, cell biology, microarrays, and protein function and analysis. [Stratagene] also provides software for pathway analysis and microarray data analysis."

29. In 2008, $2.3 billion of Agilent's net revenue was attributed to its Bio-Analytical Measurement business group, which provides "instruments, software, consumables and services that enable customers to identify, quantify and analyze the physical and biological properties of substances and products." In 2009, $2.1 billion of Agilent's net revenue was attributed to its Bio-Analytical Measurement business group. In 2010, $1.5 billion of Agilent's net revenue was attributed to its Life Science business group, which includes "liquid chromatography, mass spectrometry, microarrays, polymerase chain reaction (PCR) instrumentation, related bioreagents, electrophoresis, laboratory automation and robotics, software and informatics, nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI) systems, and related

consumables and services." Many of the Agilent products responsible for these revenues were utilized by Agilent customers with Agilent's knowledge and encouragement to practice methods that infringed upon the '179 Patent.

30. By way of example only, and upon information and belief, the Agilent 2100 Bioanalyzer product has been used by Agilent customers to determine a single nucleotide polymorphism, specifically the G20210A mutation, which is in a non-coding region of the multi-allelic prothrombin gene and results in a hereditary predisposition to venous thrombosis. In practice, Agilent customers have amplified DNA using PCR and analyzed the amplified DNA using RFLP and the 2100 Bioanalyzer to determine the presence of the G20210A mutation and a predisposition to venous thrombosis. These activities directly infringed the '179 Patent. Upon information and belief, Agilent encouraged these activities and knew that the activities by its customers infringed the '179 Patent.

31. Also, by way of example only, and upon information and belief, Agilent's GeneSpring Analysis Platforms have been used by Agilent customers to infringe at least one claim of the '179 patent when used to analyze DNA in the non-coding region. Indeed, Agilent's GeneSpring GT marketing materials state that the platform "enables you to quickly identify genotype-phenotype relationships using a comprehensive set of linkage and association algorithms." Upon information and belief, GeneSpring GT has been used by Agilent customers to identify non-coding markers which provide useful information concerning linked DNA. Agilent's customers have then used this information to analyze DNA for such non-coding markers using amplification. These activities directly infringed the '179 Patent. Upon

information and belief, Agilent encouraged these activities and knew that the activities by its customers infringed the '179 Patent.

32. Also, by way of example only, Agilent has sold a range of Comparative Genomic Hybridization Microarray products that have been used by Agilent customers to infringe at least one claim of the '179 Patent. These microarrays enable characterization of genetic variations, including those caused by developmental abnormalities, disease susceptibility and differential drug responses. For example and upon information and belief, Agilent has sold HD-CGH Custom Arrays that were developed by Agilent specifically for Agilent customers and enabled amplification of genes using PCR and contained probes including non-coding sequences on chromosomes to be analyzed to determine the presence or absence of non-coding sequences associated with multi-allelic genes. These activities infringed one or more claims of the '179 Patent. Upon information and belief, Agilent encouraged these activities and knew that these activities by its customers infringed the '179 Patent.

33. Agilent had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities. In addition, 'Agilent has had actual knowledge of the '179 Patent since at least as early as June 2006 when GTG contacted Agilent to offer it a license to the '179 Patent. Rather than negotiate a license to the '179 Patent with GTG at that time, Agilent took unreasonable positions as to why the use of its products did not infringe the '179 Patent, declined GTG's offered license, and continued to sell products that enabled and encouraged its customers to directly infringe the '179 Patent.

B. BRISTOL-MYERS

34. Bristol-Myers is actively engaged in pharmacogenetics. Bristol-Myers claims that it: "engages in genomic research for the purpose of improving human health care, and we endorse an open and informed public dialogue on all aspects of genomic research. . . . At Bristol-Myers Squibb, we currently use biomarkers in most of our first-in-human studies. Pharmacogenomic studies, which combine traditional pharmaceutical sciences such as biochemistry with annotated knowledge of genes, proteins and single nucleotide polymorphisms, are run in nearly all our oncology projects to more precisely target patient populations. . . . Bristol-Myers Squibb's bioinformatics group exploits and integrates all available sources of genome data to understand disease, for use throughout the drug discovery process."

35. Bristol-Myers is a member of the Industry Pharmacogenomics Working Group, a voluntary and informal association of pharmaceutical companies engaged in research in the science of pharmacogenomics. Bristol-Myers is also a member of The SNP Consortium (the "TSC"). According to TSC marketing materials: "The goal of the TSC allele frequency/genotype project is to determine the frequency of certain SNPs in three major world populations . . . to enhance the understanding of disease processes and facilitate the discovery and development of safer and more effective medications." Given that the majority of the human genome is covered by the non-coding region, many of these SNPs are located in non-coding regions.

36. Bristol-Myers has also actively advocated pharmacogenetic research and has undertaken its own studies and has supported third party studies via grants and funding. These studies support Bristol-Myers' investigations of gene polymorphisms associated with disease and drug response.

37. Bristol-Myers' pharmacogenetic activities include the analysis of non-coding DNA markers. By way of example only, and upon information and belief, Bristol-Myers has analyzed (or directed others to analyze) non-coding markers in connection with its Coumadin (Warfarin) and Plavix (Clopidogral) drugs using methods that infringe the '179 Patent. These Bristol-Myers products received label changes to reflect pharmacogenomic information in 2007 and 2009, respectively.

38. Warfarin is the most widely prescribed anticoagulant for thromboembolic therapy in North America and Europe. Approximately 2 million people are initiated on Warfarin therapy each year to prevent blood clots, heart attacks and stroke. However, Warfarin dosage requirements are highly variable both inter-ethnically and inter-individually. Mutations in vitamin K epoxide reductase complex, subunit 1 (VKORC1) have been identified to be associated with Warfarin effect and dosage. According to the FDA's adverse events reporting database: "complications from Warfarin (Bristol-Myer's Squibb's Coumadin) are the second most common reason for patients to go to the emergency room. . . ." Thus the use of pharmacogenetic testing is very important to the safety and efficacy of Coumadin. Coumadin is a commercially available form of Warfarin registered to Bristol-Myers Squibb Pharma Company, a division of Bristol-Myers. There is currently, and has been for some time, no direct competitor drug to Coumadin on the market.

39. A polymorphism in the promoter region of VKORC1 gene, -1639 A>G, is highly associated with inter-individual variability in Warfarin dose requirements. Promoter polymorphism -1639 A>G is listed by the FDA as a bio-marker for pharmacogenomic testing for

Warfarin sensitivity. It is claimed that 30% of dosage variability in Caucasian population is attributed to the VKORC1 gene alone.

40. In 2007, Bristol Myers began including information about the importance of -1639 A>G and Warfarin dosage in the package insert of Coumadin. The importance of the VKORC1 genotype in deciding the initial dosage is also emphasized in the inserts. In 2009, the International Warfarin Pharmacogenetics Consortium (IWPC) developed a Warfarin dose prediction algorithm using findings from nine different countries and based on the relationship between dose requirements and the known clinical and genetic factors.

41. There are four FDA approved tests (Nanosphere, Autogenomics, ParagonDX and Osmetech) available for Warfarin testing. In addition, numerous laboratory developed test are also available. The first genetic test to be cleared by the FDA for Warfarin resistance testing was Nanosphere's genetic test. The test was cleared in September 2007, shortly following Bristol-Myers' labeling change for Coumadin.

42. Upon information and belief, Bristol-Myers has (and/or has direct others to) amplify DNA with a primer pair spanning a DNA sequence containing the -1639A>G polymorphism which Bristol-Myers has associated with inter-individual variability in Warfarin dosage requirements. The analysis of this non-coding DNA marker in this manner thus infringed upon claims of the '179 Patent.

43. With respect to Clopidogral, its metabolism is enhanced by CYP2C19*17 promoter polymorphism -806 C>T. Bristol-Myers has funded and/or provided assistance to several studies to investigate the impact of genetic polymorphisms on metabolism of

Clopidogral, with at least two such studies analyzing the non-coding CYP2C19*17 promoter polymorphism.

44. Plavix, Bristol-Myers' commercially available form of Clopidogral, is Bristol-Myers' largest selling drug, accounting for $6.7 billion sales in 2010. This drug is marketed by Bristol-Myers along with Sanofi-Aventis.

45. In March 2009, the FDA announced a new Plavix warning alerting doctors and patients to the effectiveness of Plavix depending on CY2C19 genotype. Enhanced metabolisers carry CYP2C19*17 allele, which is a promoter polymorphism (-806 C>T). This non-coding promoter polymorphism is not included in the Plavix package insert, however it is included in the genetic tests offered by laboratories to test effectiveness of Clopidogral.

46. Upon information and belief, Bristol-Myers has (and/or has direct others to) amplify DNA with a primer pair spanning a DNA sequence containing the -806 C>T polymorphism which Bristol-Myers has associated with metabolism of Clopidogral. The analysis of this non-coding DNA marker in this manner thus infringed upon claims of the '179 Patent.

47. Bristol-Myers has also performed (and/or directed others to perform) other genotyping activities that infringed one or more claims of the '179 Patent. Upon information and belief, these activities were also directly related to the safety and efficacy and thus the development and/or sale of various drugs.

48. By way of example only, Bristol-Myers conducted a study in 2008 (by Ranade et al.), to study the role of genetic variation associated with highly active antiretroviral therapy. The study investigated almost 300 SNPs in 135 candidate genes. The study genotyped "the

entire cohort for eight SNPs in resistin, including four newly identified by sequencing and two from phylogenetically conserved regions." Some of these SNPS were in non-coding regions. It was determined in the study that "none was as significantly associated as the SNP in intron 2. . . . Furthermore, haplotype analysis revealed that only haplotypes bearing the intron 2 SNP, rs3219177, were significantly associated with cluster membership. . . . Taken together these results indicate that this SNP is potentially causative." This study was related to HIV treatment. Bristol-Myers claims that it: "has long been a leader in developing innovative HIV/AIDS medications and treatments and will continue to make HIV/AIDS research a top priority. . . . Our company allocates substantial R&D resources to developing new medicines and treatments in the global fight against HIV/AIDS. We are exploring new ways to attack the AIDS virus and new ways to help make treatments for patients to take."

49. Also by way of example only, Bristol-Myers conducted a study in 2007 (by Zhang et al.), involving PCR amplification of the UGTA1*28 polymorphism, to study characterization of the UDP glucuronosyl transferase activity of human liver microsomes genotyped for the UGT1A1*28 polymorphism.

50. Also by way of example only, Bristol-Myers funded a study in 2006 (by Florez et al.), concerning the association of two non-coding SNPs rs12255372 and rs7903146 of the TCF7L2 gene with the progression to diabetes. Genotyping was performed in the study using allele specific primers.

51. Upon information and belief, Bristol-Myers had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities.

C. ESTA

52. ESTA has facilities in Colorado and California and offers seed genetics, seed quality and plant health services to others. According to its literature, ESTA's "high-throughput DNA laboratory is located in [the] Colorado facility and is supported by a network of Eurofins DNA laboratories throughout the world."

53. ESTA genetic services consist of two divisions: "the HTP Lab which specializes in isoelectric focusing (IEF) of plant proteins for hybrid purity, varietal identification and uniformity and the DNA Lab which works with various types of DNA specific markers which encompasses everything from database applications to marker assisted selection, to screening with trait linked markers, GMO screening and more. . . . If a given hybrid turns out to be non-informative or monomorphic regarding the protein banding patterns we can then pass it on to our DNA lab which has about a 99% success rate in identifying polymorphisms in most vegetable species." ESTA's DNA Lab offers marker assisted backcrossing (MAB), marker assisted selection (MAS), database genotyping (DB), trait linked markers (TLM), and quantitative trait loci (QTL) identification services.

54. ESTA's marketing materials discuss commonly used molecular markers in the plant breeding industry. Among many desirable qualities for a marker, the top two desirable qualities are polymorphic and multi-allelic. The two markers identified by ESTA as satisfying these criteria are PCR-based Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphism (SNPs).

55. The majority of SSRs are non-coding as a majority of the genome is non-coding.

56. The majority of SNPs will also be in a non-coding DNA region. Indeed, ESTA admits that: "SNPs are abundant in all genomes and can be found approximately every kilobase (1,000 base pairs). They are also spread evenly throughout the genome. This offers the potential for generating very high density genetic maps, which will be extremely useful for developing haplotyping systems for genes or regions of interest. They may be the polymorphisms associated with the gene of interest under study and therefore direct selection of the gene is possible."

57. ESTA admits that it has "identified an optimal number of markers spaced throughout a plant's genome and studied computer simulations to determine the number of plants to test each generation. In each backcross generation, the population has a wide range for percentage of recurrent parent genome. Applying DNA markers effectively and efficiently allows you to identify the individuals that inherit the highest percentage of the recurrent parent genome plus the trait of interest. As a complement after backcrossing, our service can select individuals homozygous for the trait of interest thereby expediting commercialization of your product."

58. Upon information and belief, ESTA has also customized marker strategies to suit customer needs and also analyzes markers to select desired traits and help design desired genetic ratio of original parent and cross as a part of its MAS services.

59. ESTA has also offered database genotyping which is based on "specific sets of molecular markers for any given species" and "use the same set of DNA markers on all entries." ESTA admits on its website that genotyping can be used for inbred line development, confirmation of pedigrees and association studies. ESTA also claims to have collected a large number of trait linked markers in many species. According to ESTA: "In some cases, we have

identified molecular markers linked to genes which promote expression of a specific trait, for example a disease resistance gene. Using the marker, we can identify plants carrying the gene, including the zygosity, even in the absence of the selection agent and at any stage of plant growth."

60. Sometimes important traits are difficult to select by phenotype due to contribution by more than one gene or failure to locate a gene that is responsible for the trait. In such cases markers are used to map and identify a locus using appropriate phenotypic, genotypic and statistical analyses. A stretch of DNA containing the gene or linked markers responsible for a quantitative trait is referred to as a QTL. Upon information and belief, ESTA has offered QTL mapping and identification services by closely working with their customers "to determine the best population type to work with, number of individuals to test, number of markers to apply, and other unique aspects of your breeding program to help ensure you accomplish your program goals." ESTA also states that its "Genetic Services can identify DNA markers that span a plants genome based on your specific population(s). When analyzed on a segregating or recombinant inbred population along with appropriate phenotypic data, genomic regions can be identified which contribute to a specific phenotype. The phenotype or trait could be a quantitative trait implying many genes could contribute to the phenotype or a qualitative trait resulting from one or two major genes." Trait specific molecular marker technology (MMT) is also used by ESTA genetic services laboratory to test hybrid purity.

61. U.S. Patent Application Publication No. U.S. 2011/0041204 A1, entitled "Methods for Enhancing the Production and Consumer Traits of Plants" ("the '204 Publication") refers to the use of ESTA services sometime before the filing date of the '204 Publication on

August 12, 2009. The '204 Publication states: "Even if an enzymatic basis for a particular mutant gene is not known, and the nucleotide sequence for the gene encoding the enzyme is not known, and is not present in the Maize Genetics and Genomics or GenBank databases, the inheritance of the gene can still be determined by those having ordinary skill in the art by following nearby molecular markers on the chromosome including the gene. . . ." The '204 Publication also discusses the use of commercially available DNA polymorphisms that were tested for trait identification for the su1, se2 and sh2 genes. The sh2 gene, which is multi-allelic, can mutate from a G nucleotide to a C, T or A nucleotide in intron 15 of the gene. When the mutation (C, T or A) is expressed in a plant, such as maize, enhanced growth characteristics occur. The '204 Publication further states that ESTA provided services and used 330 SSR markers for the study described in the '204 Publication. All of those markers are, upon information and belief, in non-coding regions.

62. As described above, ESTA has utilized many non-coding DNA markers using amplified DNA with a primer pair spanning a DNA sequence containing these non-coding markers. For example, ESTA's marketing materials describe the laboratory process for seed sampling to include PCR amplification. ESTA has associated these markers with genes responsible for desired traits of economic importance. ESTA's analysis of these non-coding DNA markers thus directly infringed upon claims of the '179 Patent.

63. Upon information and belief, ESTA had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities.

D. GENESEEK

64. GeneSeek is a commercial agricultural genetics laboratory, providing agri-genomic services to research and development and commercial industries. These services include genotyping and gene mapping from sample processing through to data analysis and evaluation, and SNP profiling, marker-assisted selection, identity and diagnostics testing. GeneSeek's marketing materials list alpaca, cattle, canine, goat, horse, swine and sheep as well as maize, rice, soybean and wheat plant species as among the species that it has conducted such testing.

65. GeneSeek maintains service testing facilities in Lincoln, Nebraska that hold a range of platform technologies used to perform its agri-genomic services. These include the Sequenom MassARRAY, Li-Cor DNA Sequence Analyzer and Illumina GoldenGate and Infinium platforms.

66. According to GeneSeek: "with our technical platforms we can handle any project, from a few SNPs (single nucleotide polymorphisms) on thousands of samples to a million SNPs on a few samples. We can process any sample type including hair, blood, tissue and more. We also provide clinical veterinary diagnostics for trait and disease markers and for pathogen detection using quantitative real-time PCR and ELISA. . . . For 12 years we have been the trusted leader in providing high-throughput SNP and microsatellite genotyping services to the agricultural research and business communities."

67. GeneSeek has performed SNP profiling on Illumina products such as:

Bovine 3K Net Merit Panel 3072 SNP panel

Bovine SNP50 BeadChip – 54,609 informative SNP markers

BovineHD – High density genotyping with greater than 777,000 SNP markers

CanineHD – High density genotyping with greater than 170,000 SNP markers

PorcineSNP60 BeadChip – 62,163 informative SNP markers

OvineSNP50 BeadChip – 54,241 informative SNP markers

EquineSNP50 BeadChip – 54,602 informative SNP markers

Maize SNP50 BeadChip – 56,110 informative SNP markers

Mouse Custom 9K BeadChip

68. SNP profiling using Illumina products detects and analyzes many non-coding DNA markers. By way of example only, the "PorcineSNP60 BeadChip features more than 62,000 evenly spaced SNPs across the entire porcine genome." Since more than 95% of the pig genome is non-coding DNA, many of the probes which target "evenly spaced polymorphic SNPs" are located in non-coding regions. The use of these Illumina products thus infringed upon claims of the '179 Patent.

69. Upon information and belief, GeneSeek utilized DNA amplification as a part of its testing services. By way of example only, GeneSeek states in its marketing materials: "GeneSeek can provide genotyping services for . . . whole genome scans or marker assisted selection . . . Genotyping from plant samples that allow high throughput genotyping at very low cost. From one or two small leaf disc punches, we are able to analyze your target of interest using standard PCR, real-time PCR, SNPs or microsatellite markers."

70. Also by way of example only, GeneSeek has performed sequencing and microsatellite analysis using the Li-Cor DNA Sequence Analyzer. This platform utilizes PCR amplification of DNA to prepare samples for analysis. Upon information and belief, GeneSeek's

sequencing and microsatellite analysis included analysis of markers in non-coding regions. Thus, these activities infringed upon claims of the '179 Patent.

71. Also, by way of example only, GeneSeek's identity and diagnostics portfolio includes "Seek-Gain: Total Gain" and "Seek-Gain: Animal Gain" tests which detect and analyze the non-coding A/G variant at position +179 in the 5' untranslated region of the CCKAR gene and C/T variation at position 576 in intron 6 of the HMGA1 gene. According to GeneSeek, the "cholecystokinin type A receptor (CCKAR) genetic test is associated with physiological control of feed intake, ~5% higher daily feed intake, 3% higher daily gain, and 3% fewer days to reach 180kg, when compared to homozygotes for the A-variant" and the "high mobility group AT-hook protein 1 (HMGA1) genetic test is associated with lean mass percentage, growth and back fat in several swine breeds. Producers can test and select animals (T-variant at position 576) which are likely to be leaner and produce offspring that are leaner." Upon information and belief, these testing services infringed upon claims of the '179 Patent.

72. Also, by way of example only, GeneSeek's "Seek-Gain: Total Gain" and "Seek-Gain: Litter Size" tests detect and analyze the non-coding C/T variant in intron 4 of the EPOR gene. According to GeneSeek, there is "a genetic variant in the swine erythropoietin receptor gene associated with uterine capacity and litter size. The favorable genetic variation has demonstrated an increase in uterine capacity as well as an increase in live births in two different swine populations at USDA – MARC, including the industry-relevant BX population. In a commercial herd, an extra pig per litter was observed when comparing boars that have two copies of the favorable EPOR marker versus boars with zero copies." Upon information and belief, these testing services infringed upon claims of the '179 Patent.

73. Upon information and belief, GeneSeek had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities.

E. GSK

74. GSK is active in the area of pharmacogenetics. In its marketing materials it admits that: "GSK invests heavily in research, including genomic and genetic research, that may lead to . . . and new methods of disease detection, prevention and treatment."

75. GSK's 2005 Annual Report outlines how pharmacogenetics plays a role in the strategies used by GSK for drug development:

> Two components are needed in the early stages of finding new medicines - targets that can be shown to affect mechanisms of important pathological processes in human disease and compounds able to modulate the behavior of specific targets. Many diseases arise through complex interactions between gene variants and environmental factors. Within GSK, Genetics Research aims to take advantage of this by identifying genes which influence common diseases with large unmet medical needs and major patient burdens. These insights help in the search for targets with known relevance to the disease, and hence a greater chance of delivering benefit to the patients. Discovery Research (DR) produces the lead compounds that may influence targets which form the basis of drug discovery efforts in GSK's Centres of Excellence for Drug Discovery (CEDDs). In 2005, DR performed over 90 million assays and provided the CEDDs with 50 high-quality new lead compounds. Investment in DR has been focused on increasing the quality and quantity of the lead compounds available.

76. GSK also reports that it is "an active member of the SNP Consortium ["TSC"], collaboration between industry and the UK's Welcome Trust, contracting with academic institutions to identify and map SNPs. The information arising from the collaboration is a valuable research tool and has been placed in the public domain without IP restrictions. GSK believes that the value of SNPs lies not so much in their identification but in their association with diseases or patient response to medicines." In turn, TSC marketing materials indicate that:

"The goal of the TSC allele frequency/genotype project is to determine the frequency of certain SNPs in three major world populations . . . to enhance the understanding of disease processes and facilitate the discovery and development of safer and more effective medications." Given that the majority of the human genome is covered by the non-coding region, many if not most of the SNPs identified and mapped by TSC are located in non-coding regions.

77. GSK has actively advocated pharmacogenetic research and has undertaken its own pharmacogenetic studies and has supported third party studies via grants and funding. These studies support GSK's investigations of gene polymorphisms associated with disease and drug response.

78. GSK's pharmacogenetic activities include the analysis of non-coding DNA markers. As further described herein, upon information and belief, GSK has analyzed non-coding markers in connection with the drugs Ziagen (Abacavir) and Tykerb (Lapatinib) using methods that infringed the '179 Patent. These activities were related to the safety and efficacy and thus the development and/or sale of Ziagen and Tykerb.

79. Abacavir is an antiviral medication used in combination with other antiretroviral drugs for the treatment of HIV-1 infection. Ziagen is a commercially available form of Abacavir that is manufactured and sold in the United States by GSK.

80. In 2005, GSK conducted a study to identify genetic markers associated with presence or absence of Abacavir hypersensitivity. Using a genome wide SNP genotyping approach that included both coding and non-coding regions, genes associated with immune response were studied including HLA and TNF α. Many TNF α variants are within non-coding DNA regions.

81. Upon information and belief, GSK has (and/or through service providers acting at GSK's direction) amplified DNA with a primer pair spanning a DNA sequence containing a non-coding polymorphism which GSK has associated with hypersensitivity to Abacavir. These activities thus directly infringed the '179 Patent. Upon information and belief, GSK encouraged these activities and knew that the activities infringed the '179 Patent.

82. Lapatinib is a tyrosine kinase inhibitor and is an oral medication for breast cancer. Tykerb is a commercially available form of Lapatinib that is manufactured and sold in the United States by GSK. It inhibits HER2 and EGFR and is used in HER2 positive women with metastatic breast cancer. In a 2007 article in Pharmaweek, GSK's Allen Roses described the use of densely mapped SNPs within cytochrome P450, genes particularly in CYP2C19, that include non-coding polymorphisms. Some variants were associated with side effects to this drug.

83. Upon information and belief, GSK has also conducted and/or funded at least two pharmacogenetic studies related to the use of Lapatinib in the treatment of breast cancer, and has directly (and/or through service providers acting at GSK's direction) amplified DNA with a primer pair spanning a DNA sequence containing a non-coding polymorphism which GSK has associated with the efficacy of Lapatinib. The analysis of this non-coding polymorphism in this manner thus infringed upon claims of the '179 Patent. These activities thus directly infringed the '179 Patent. Upon information and belief, GSK encouraged these activities and knew that the activities infringed the '179 Patent.

84. GSK has also performed (and/or directed others to perform) other genotyping activities that infringed one or more claims of the '179 Patent. Upon information and belief,

these activities were also directly related to the safety and efficacy and thus the development and/or sale of various GSK drugs.

85. By way of example only, GSK conducted a study in 2009 that was reported in an article entitled "CTLA4 Gene Polymorphisms are Associated with Chronic Bronchitis" in the European Respiratory Journal. It was reported in the article that:

> Six CTLA4 SNPs were significantly associated with chronic bronchitis in the ICGN cohort ($0.0079 \leq p \leq 0.0432$) with three being replicated with the same directionality of association in the Bergen cohort ($0.0325 \leq p \leq 0.0408$). One of these replicated SNPs (rs231775) encodes the Thr to Ala substitution at amino acid position 17. Haplotype analyses supported the results of single SNP analyses . . .
>
> As CTLA4 is located near the region on chromosome 2q that showed significant linkage with COPD-related phenotypes both in the Boston Early-Onset COPD Study 9-11, as well as in general population pedigrees 12, we hypothesized that CTLA4 single nucleotide polymorphisms (SNPs) would be associated with COPD and COPD-related phenotypes, including the severity of airflow limitation and the presence of chronic bronchitis . . .
>
> Genotyping in the two cohorts was performed with the Illumina array-based custom SNP genotyping platform.
>
> We found significant associations with the other COPD-related phenotype investigated, chronic bronchitis. In the ICGN families, six out of the nine CTLA4 SNPs genotyped (SNPs rs926169, rs11571316, rs231775, rs231779, rs3084243 and rs231725) were significantly associated with the chronic bronchitis phenotype ($0.0079 \leq p \leq 0.0432$). Associations with SNPs rs231775 and rs3087243 ($p = 0.0081$ and 0.0079, respectively) were significant even after a correction for multiple testing. We replicated the significant association with chronic bronchiitic phenotype in the COPD cases of the Bergen cohort for four of the six CTLA4 SNPs (rs11571316, rs231775, rs3087243 and rs231725; $0.0325 \leq p \leq 0.0408$) identified in the ICGN study. After evaluating the risk allele, we found that three of the replicated significant SNPs (rs231775, rs3087243 and rs231725) for chronic bronchitis have the same directionality of association in the two populations.
>
> We did not detect any significant association with chronic bronchitis in control subjects without COPD in the Bergen cohort, suggesting that CTLA4 is associated with chronic bronchitis among COPD subjects . . .

> In the ICGN study, we identified two haplotype blocks. Several of the significantly associated SNPs (rs926169, rs11571316, rs231775 and rs231777) were located in block 1, while three other significantly associated SNPs (rs231779, rs3087243 and rs231725) were located in block 2.

Among the nine SNPs genotyped in the study, 8 are located in non-coding regions, namely rs926169 (5' intergenic region upstream of CTL4 gene), rs733618, rs11571316, rs16840252 located in the promoter region, rs231777, rs231779 located in the intron, rs3087243, and rs231725 located downstream of the CTL4 gene. Upon information and belief, GSK has directly (and/or through service providers acting at GSK's direction) analyzed these SNPs in a manner that infringed upon claims of the '179 Patent. Upon information and belief, GSK encouraged these activities and knew that the activities infringed the '179 Patent.

86. Also by way of example only, GSK conducted a study in 2007 that was reported in an article entitled "Association of PTGDR Gene Polymorphisms with Asthma in Two Caucasian Populations" in the Genes & Immunity Journal. Therein it was reported that:

> The prostanoid DP receptor (PTGDR) is shown to be involved in the asthma patho-physiology and the results from the published genetic association studies are inconsistent. Four single nucleotide polymorphisms (SNPs) in PTGDR were genotyped in 342 and 294 families from UK and Denmark respectively. Asthma and asthma-related phenotypes were analyzed using family-based association analyses. In the UK families, a promoter polymorphism (-731A/G) showed significant associations with asthma (P=0.0022), atopic asthma (P=0.0044), bronchial hyper reactivity or BHR (P=0.00120) and strict asthma (P=0.0008). The p-values for asthma, BHR and strict asthma were significant even after the most stringent correction for the number of markers and the number of phenotypes analyzed (<0.0031). An intronic polymorphism (+6651C/T) also showed significant associations with asthma (P=0.0302) atopic asthma (P=0.0131), BHR (P=0.0249) and strict asthma (P=0.0261). In the Danish families, an intronic asthma (P=0.0348), BHR (P=0.0033) and strict asthma (P=0.0381). The results of haplotype analyses supported the ones of the single SNP analyses. Thus, we demonstrated significant evidence of association between polymorphisms in PTGDR with asthma phenotypes in the two Caucasian populations.

Upon information and belief, GSK has directly (and/or through service providers acting at GSK's direction) analyzed non-coding polymorphisms in PTGDR in a manner that infringed upon claims of the '179 Patent. Upon information and belief, GSK knew that these activities infringed the '179 Patent.

87. Upon information and belief, GSK had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities. In addition, GSK has had actual knowledge of the '179 Patent since at least as early as November 2001 when GTG contacted GSK regarding licensing of the '179 Patent and other patents.

F. HOLOGIC

88. Upon information and belief, in July of 2008 Hologic acquired Third Wave Technologies, which "develops and markets molecular diagnostic reagents for a wide variety of DNA and RNA analysis applications based on its proprietary Invader chemistry." In 2008, nine weeks of sales attributed to Third Wave diagnostic products was approximately $5.9 million. In 2009, revenue attributed to Third Wave diagnostic products was approximately $37.1 million. One of these diagnostic product lines is sold under the name Invader.

89. Upon information and belief, the Invader chemistry involves two simultaneous reactions. In the primary reaction, a probe and an Invader oligonucleotide anneal to a DNA target sequence of interest. If a mutation or variant is present, a one base overlapping nucleotide structure is cleaved and released as a flap. This reaction is repeated such that the number of flaps released is related to the amount of the target sequence of interest in the sample, which offers a quantitative detection of the genes present in the sample. In the secondary reaction, the flaps

attach to a fluorescent resonance energy transfer cassette. Once the flap attaches and the one base overlapping nucleotide is recognized, it releases a fluorescent signal that can be detected with a multi-well fluorometer. The Invader technology has been used to infringe at least one claim of the '179 Patent.

90. The Invader chemistry is used in several products offered by Hologic, including InPlex CF, Life Science Research Kit, Chromosome Specific Panels, Invader Whole Genome Screening Panels, Low Density Genome-wide Panel, Higher Density SNP Panel, Invader UGT1A1 Molecular Assay, and through Hologic AgBio services, which include SNP genotyping.

91. By way of example only, the Invader chemistry used in the InPlex CF test kits detects and identifies mutations and variants in the cystic fibrosis transmembrane conductance regulator gene, including mutations and variants recommended for detection by the 2004 American College of Medical Genetics. Many of these mutations and variants are found only in non-coding DNA regions.

92. Hologic AgBio has also offered services using InvaderPlus technology, including SNP genotyping. Hologic has also offered the Universal Invader product, which allows researchers to design reactions that use the Invader chemistry. Both InvaderPLUS and Universal Invader involve PCR amplification followed by the Invader chemistry reaction and can be used and, upon information and belief, have been used to conduct methods covered by the '179 Patent.

93. Upon information and belief, Hologic had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities.

G. MERIAL

94. Merial has offered a series of genetic tests for DNA markers associated with quality traits in beef and dairy cattle under the trade name "Igenity." According to Merial's marketing materials, "Igenity helps you understand and manage the potential for animals to perform and transmit traits of economic importance." The Igenity DNA marker profile includes numerous non-coding DNA markers associated with genes responsible for these traits. Merial's analysis of these non-coding DNA markers either directly and/or through service providers acting at the direction of Merial infringed upon claims of the '179 Patent.

95. By way of example only, the Igenity TenderGENE tests for three DNA markers including two SNPs in the calpain gene (CAPN 316 and CAPN 4751) and one in the calpastatin gene (UoG-CAST 1). These genetic markers are associated with meat tenderness. At least the CAPN 4751 and UoG-CAST 1 markers are located solely in non-coding regions. Upon information and belief, Merial directly and/or through service providers acting at the direction of Merial amplified DNA with a primer pair spanning a DNA sequence containing these non-coding markers. Merial's analysis of these markers thus infringed upon claims of the '179 Patent. Upon information and belief, Merial encouraged these activities and knew that the activities infringed the '179 Patent.

96. Also, by way of example only, Merial has tested for markers associated with the leptin gene, a gene associated with appetite and metabolism and the propensity for meat to marble. The analysis included several non-coding DNA markers. More specifically, Merial's International Patent Application No. WO 2006/096427 A2 ("the '427 Publication") relates to identification of leptin gene polymorphisms and their use in determining genotype/phenotype

association in live stock. Three polymorphisms, UASMS 1, UASMS 2, and UASMS 3, are identified in the '427 Publication as being located in promoter or non-coding DNA regions. The '427 Publication also describes the primer pair sequences for the amplification of DNA sequences containing these non-coding markers. Merial's analysis of these markers either directly and/or through service providers acting at the direction of Merial infringed upon claims of the '179 Patent. Upon information and belief, Merial encouraged these activities and knew that the activities infringed the '179 Patent.

97. Also, by way of example only, in the publication entitled "Polymorphisms In Two Positional Candidate Genes In The Bovine Chromosome 14 Are Associated With Carcass Merit In Beef Cattle," presented at the Plant & Animal Genomes XV Conference held January 13-17, 2007 in San Diego, California, Merial states:

> Nine unique polymorphism were identified in DECR1, including 4 exonic SNPs and 4 unique polymorphisms were identified in the CBFA2T1 gene, all of which are intronic. . . . Two of the exonic polymorphisms produced amino acid composition changes from isoleucine to valine and valine to methionine. The latter appears to be located in a conserved region where important catalytic reactions occur.

In the publication entitled "Polymorphisms In The Bovine Fibroblast Growth Factor 8 (FGF8) Gene Are Associated With Carcass Quality And Growth Traits In Beef Cattle," presented at the Plant & Animal Genomes XV Conference held January 13-17, 2007 in San Diego, California, Merial states:

> In this experiment, 464 animals from an experimental composite line comprised of Angus, Charolais and Hybrid steers were sequenced for FGF8 introns and exons (NW_930497.1). Four unique polymorphisms were identified, two exonic and two intronic. Single locus and haplotype association analysis . . . were carried out resulting in a number of significant associations with carcass quality and growth traits in beef cattle; including carcass lean meat yield %, ultrasound marbling and birth weight.

In the publication entitled "Genetic And Phenotypic Relationships Of Serum Leptin Concentration With Performance, Feed Efficiency, And Carcass Merit Of Feedlot Cattle," J. Animal Science, pp 1-30, Nkrumah JD et al. (2007), Merial states:

> Mutations in the leptin gene or its promoter are associated with differences in serum leptin concentrations and other economically relevant traits in beef and dairy cattle. Indeed, polymorphisms in the bovine leptin promoter have been shown to have strong associations with serum leptin concentrations as well as body fatness.

98. Upon information and belief, Merial has either directly and/or through service providers acting at the direction of Merial analyzed the non-coding markers identified in these publications and thus infringed upon claims of the '179 Patent. Upon information and belief, Merial encouraged these activities and knew that the activities infringed the '179 Patent.

99. Since the first offering of its Ingenity tests through the expiration of the '179 Patent, Merial has added numerous DNA markers to its Igenity profile. Upon information and belief, many of those markers are located in non-coding DNA regions. Upon information and belief, Merial has either directly and/or through service providers acting at the direction of Merial used DNA amplification to analyze these non-coding DNA markers and to detect the genes responsible for the associated traits of economic importance. These activities thus infringed upon claims of the '179 Patent. Upon information and belief, Merial encouraged these activities and knew that the activities infringed the '179 Patent.

100. Upon information and belief, Merial had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities.

H. PFIZER

101. Pfizer is active in the area of pharmacogenetics. Pfizer acknowledges that the use of pharmacogenetic information can alter the course of its drug development or change the drug labeling for on-shelf pharmaceutical drugs in order to retain FDA approval.

102. Pfizer has actively advocated pharmacogenetic research, both in terms of therapeutic effect as well as adverse effects. Pfizer has also undertaken numerous pharmacogenetic studies (clinical trials) and funds studies to investigate gene polymorphisms associated with disease and drug response.

103. Pfizer's pharmacogenetic activities include the analysis of non-coding DNA markers. As further described below, and by way of example only and upon information and belief, Pfizer has either directly and/or through others acting at the direction of Pfizer analyzed non-coding DNA markers in connection with the drugs Camptosar (Irinotecan) and Zoloft (Sertraline) using methods that infringed claims of the '179 Patent. Upon information and belief, Pfizer encouraged these activities and knew that the activities infringed the '179 Patent.

104. With respect to Irinotecan, Pfizer directly (and/or through service providers acting at Pfizer's direction) performed genotyping activities in 2005 as well as funded numerous studies undertaken by third parties using the inventions of the '179 Patent. The genotyping test that is related to the drug's metabolism and thus toxicity interrogates a non-coding polymorphism of the multi-allelic UGT1A1 genetic locus. These activities are directly related to the safety of use of Irinotecan and therefore sale of the drug.

105. Irinotecan hydrochloride was originally developed in Japan by the Yakult Honsha Company. Licensing rights for clinical development in the U.S. were granted to Pharmacia,

whereas similar rights in Europe were granted to Aventis. Irinotecan was first approved in the U.S. for the treatment of metastatic colorectal cancer after failure of first-line treatment with 5-FU. This initial approval was based on tumor response rate data from phase II, uncontrolled studies. Conditional marketing authorization in the U.S. was granted in 1996 under FDA regulations designed to accelerate approval of new and promising drugs for serious or life-threatening illnesses. Subsequently, Aventis completed two European randomized, phase III studies comparing second-line Irinotecan therapy with best supportive care or with infusional 5-FU-based therapy and provided the data from these trials to Pharmacia. The survival advantages associated with Irinotecan use in each of these trials was the basis for full FDA approval for Irinotecan as second-line therapy for patients with metastatic colorectal cancer in September 1998. It was approved as a first-line therapy in April 2000.

106. Pfizer acquired Pharmacia in 2003 for $60 Billion comprising assets related to Irinotecan in the U.S. and others. Pfizer then acquired the European rights for Irinotecan from Aventis in 2004 for $620 Million. Prior to patent expiration, Irinotecan was Pfizer's major cancer treatment drug with global revenues of $969 Million in 2007. Campotosar is a commercially available form of Irinotecan that is manufactured and sold in the United States by Pfizer. Pfizer lost exclusivity for Irinotecan in February 2008 and July 2009, in U.S. and Europe, respectively. Currently, Pfizer's global revenue from Irinotecan is $117 Million.

107. The UGT1A1*28 allele is a promoter polymorphism reported to effect drug efficacy of Irinotecan. Pfizer collaborated in a clinical trial entitled "Toxicity/Benefit Ratio Optimization of Chemotherapy in Colorectal Cancer (CRC) Patients by Determination of Individual Genotypic Determinants" in June 2005 and completed the trial in December 2008.

The study involved genotyping patients for the UGT1A1 polymorphisms prior to first administration of Irinotecan.

108. The FDA recommended label changes to Irinotecan to include UGT1A1 genetic information in July 2005. In conjunction with the label changes, the FDA simultaneously approved the first commercial kit offered by Third Wave Technologies for genotyping the UGT1A1 polymorphism. Subsequently, there was an increase in both the number of laboratories offering UGT1A1 genotyping as well as availability of related reagents and kits.

109. Pfizer has encouraged in prescribing information that "UGT1A1 activity is reduced in individuals with genetic polymorphisms that lead to reduced enzyme activity such as the UGT1A1*28 polymorphism. Approximately 10% of the North American population is homozygous for the UGT1A1*28 allele. In a prospective study, in which Irinotecan was administered as a single-agent on a once-every-3-week schedule, patients who were homozygous for UGT1A1*28 had a higher exposure to SN-38 than patients with the wild-type UGT1A1 allele. . . . Individuals who are homozygous for the UGT1A1*28 allele are at increased risk for neutropenia following initiation of CAMPTOSAR treatment. A reduced initial dose should be considered for patients known to be homozygous for the UGT1A1*28 allele." Acting upon these directions by Pfizer, laboratories and doctors conducting analysis of DNA for the UGT1A1*28 allele in connection with the prescribing of Camptosar directly infringed upon claims of the '179 Patent. Upon information and belief, Pfizer encouraged these activities and knew that the activities infringed the '179 Patent. Upon information and belief, Pfizer has amplified DNA with a primer pair spanning a DNA sequence containing the UGT1A1*28

polymorphism which Pfizer has associated with increased risk for neutropenia. These activities by Pfizer directly infringed upon claims of the '179 Patent.

110. With respect to Sertraline, Pfizer has also performed genotyping activities that relate to the inventions of the '179 Patent. The genotyping test that is related to the drug's metabolism and thus toxicity interrogates a non-coding polymorphism (5HTTLPR) of the multi allelic SLC6A4 genetic locus. These activities are directly related to the safety of use of Sertraline and therefore sale of the drug.

111. Zoloft is a commercially available form of Sertraline that is manufactured and sold in the United States by Pfizer. Zoloft was one of Pfizer's blockbuster drugs generating $2.11 Billion in 2006. It is used to treat major depressive disorder, panic disorder, obsessive-compulsive disorder (OCD) in adults and children, post-traumatic stress disorder (PTSD), premenstrual dysphoric disorder (PMDD) and social anxiety disorder (SAD). Zoloft is approved for acute and long-term use in all of these indications, with the exception of PMDD. The drug went off patent in June 2006.

112. In 2006, Pfizer's sister company Pfizer Australia funded a study entitled "Serotonin Transporter Polymorphisms and Clinical Response to Sertraline Across Ethnicities," which investigated the "relationship between clinical response, adverse effects, Sertraline (SERT) plasma concentrations and the genetic polymorphism of the serotonin transporter gene-linked polymorphic region (5HTTLPR)." Upon information and belief, Pfizer has directly (and/or through service providers acting at Pfizer's direction) replicated and/or conducted further analysis consistent with the 2006 study and in doing so has amplified DNA with a primer pair spanning a DNA sequence containing the 5-HTTLPR polymorphism which Pfizer has associated

with genes effecting serotonin responsiveness. The analysis of this non-coding DNA marker in this manner thus directly infringed upon claims of the '179 Patent. Upon information and belief, Pfizer encouraged these activities and knew that the activities infringed the '179 Patent.

113. Upon information and belief, and as described below, Pfizer has also performed other genotyping activities that infringed one or more claims of the '179 Patent. These activities were also directly related to the safety and efficacy and thus the development and/or sale of various drugs.

114. By way of example only, Pfizer's International Patent Application No. WO2005/090601 and entitled "Biomarker of Hypertension" ("the '601 Publication") describes the analysis of the -54E3 polymorphism of the GUCY1 A2 gene. The -54E3 polymorphism is an A/G variant which occurs in intron 2 of the GUCY1A2 gene. The '601 Publication describes the use of PCR and other amplification techniques for amplifying genomic DNA samples and analysis of those samples using a number of tools. For example, Claim 5 of the '601 Publication reads:

> A kit for the diagnosis of hypertension or a predisposition to hypertension in a human subject, or for the selection of a human subject likely to respond to treatment with an sGC activator compound, comprising means for identifying the genotype of the -54E3 A/G single nucleotide polymorphism of the human GUCY1 A2 gene, and/or one or more single nucleotide polymorphisms of GUCY1A2 which is/are in linkage disequilibrium therewith, in DNA taken from the subject.

Upon information and belief, Pfizer has replicated and/or conducted further analysis consistent with the '601 Publication directly (and/or through service providers acting at Pfizer's direction) and has thus infringed upon claims of the '179 Patent. Upon information and belief, Pfizer encouraged these activities and knew that the activities infringed the '179 Patent.

115. Also by way of example only, Pfizer's International Publication No. WO 2005/031341 and entitled "Methods for Predicting Development of Insulin Resistance" ("the '341 Publication") describes the analysis of the -359 polymorphism of the P110β gene using PCR amplification techniques for amplifying genomic DNA samples. In addition, the '341 Publication describes: "In one aspect of any of the methods of the invention, the step of determining whether the DNA of subject comprises a particular P110β allele can be performed using a nucleic acid molecule that specifically binds a P110β nucleic acid molecule. Preferably, the P110β allele comprises detecting a polymorphism in a transcriptional regulatory region, in further preferred aspects, the methods of the invention comprise determining whether the DNA of an individual comprises a T or a C at position 100 of SEQ ID NO 1, or at position -359 (position 359 upstream from the start codon) of the P110β gene. This may thus comprise determining whether the genomic DNA of an individual comprises a P110β allele, whether mRNA obtained from an individual comprises a P110β allele." Upon information and belief, Pfizer has replicated and/or conducted further analysis consistent with the '341 Publication directly (and/or through service providers acting at Pfizer's direction) and has thus infringed upon claims of the '179 Patent. Upon information and belief, Pfizer encouraged these activities and knew that the activities infringed the '179 Patent.

116. Upon information and belief, Pfizer had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities. In addition, Pfizer has had actual notice of the '179 Patent since at least as early as March 2008 when its business unit, Pfizer Animal Health, acquired Catapult Genetics Pty., Ltd. and Bovigen, LLC who are both licensees of the '179 Patent. During the acquisition

process for these GTG licensees, there were numerous communications between Pfizer and GTG concerning the licenses to the '179 Patent.

I. 454

117. In January of 2007, 454 began offering for sale the Genomic Sequencer FLX system ("GS FLX"). 454 was acquired by Roche Diagnostics ("Roche") in March of 2007. Upon information and belief, 454 began using and testing the Genomic Sequencer Junior system ("GS Junior") at least as early as 2009. Both the GS FLX and GS Junior systems use a range of kits containing reagents provided by Roche. The use of these systems has infringed at least one claim of the '179 Patent.

118. The GS FLX and GS Junior systems are used to analyze genomic DNA and PCR products. These materials are segmented into 300-800 base pair fragments if necessary. The fragments are immobilized to a DNA capture bead. The fragmented sequence containing capture bead can be amplified using emulsion PCR and analyzed using the systems.

119. By way of example only, in a 2009 study entitled "Rare Variants of IFIH1, a Gene Implicated in Antiviral Responses, Protect against Type 1 Diabetes," sequencing was performed by scientists, including a 454 employee named David Riches, using the GS FLX system to determine if there was a genetic variation linked to Type 1 Diabetes ("T1D"). The GS FLX system was used to re-sequence genes believed to be associated with T1D. The study analyzed both coding and non-coding regions of the IFIH1 gene, which is multi-allelic. Thus, when used to analyze the non-coding region of the multi-allelic IFIH1 gene by the process described above, use of the GS FLX system infringed at least one claim of the '179 Patent. Upon

information and belief, 454 directed these activities and/or encouraged these activities and knew that the activities infringed the '179 Patent.

120. Upon information and belief, 454 had actual knowledge of the '179 Patent during times relevant to this action through at least its research, development and/or patent application activities. In addition, 454 has had actual knowledge of the '179 Patent since at least as early as August 2006 when 454 and GTG engaged in licensing discussions regarding the '179 Patent and other patents.

V. FIRST CLAIM FOR RELIEF
(Patent Infringement by Agilent – U.S. Patent No. 5,612,179)

121. GTG incorporates by reference each and every allegation in paragraphs 1 through 120 as though fully set forth herein.

122. Agilent has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products, the use of which by Agilent's customers has directly infringed one or more claims of the '179 Patent. Upon information and belief, Agilent was aware of and encouraged its customers to use Agilent products in a manner that infringed the '179 Patent, and intended to cause that infringement. Alternatively, Agilent willfully blinded itself to the infringing nature of its customer's use of Agilent products to infringe the '179 Patent. Agilent has thus induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b).

123. Agilent has contributed to direct infringement of the '179 Patent by others by offering to sell, selling or providing one or more items which constitute a material part of inventions defined by claims of the '179 Patent, including, without limitation, Agilent's HD-CGH

Custom Arrays, knowing the same to have been especially made or adapted for use in an infringement of the '179 Patent, which components are not staple articles or commodities of commerce suitable for substantial non-infringing use in violation of 35 U.S.C. § 271(c).

124. Upon information and belief, Agilent's infringement of the '179 Patent has been willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

125. GTG has been damaged as a result of Agilent's infringing conduct. Agilent is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

VI. SECOND CLAIM FOR RELIEF
(Patent Infringement by Bristol-Myers – U.S. Patent No. 5,612,179)

126. GTG incorporates by reference each and every allegation in paragraphs 1 through 125 as though fully set forth herein.

127. Bristol-Myers has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products and/or services that infringed one or more claims of the '179 Patent in violation of 35 U.S.C. § 271(a) and/or has induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b) and/or has contributed to direct infringement of the '179 Patent by others by offering to sell, selling or providing one or more items which constitute a material part of inventions defined by claims of the '179 Patent, knowing the same to have been especially made or adapted for use in an infringement of the '179 Patent, which components are not staple

articles or commodities of commerce suitable for substantial non-infringing use in violation of 35 U.S.C. § 271(c).

128. Bristol-Myers' actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

129. GTG has been damaged as a result of Bristol-Myers' infringing conduct. Bristol-Myers is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

VII. THIRD CLAIM FOR RELIEF
(Patent Infringement by ESTA – U.S. Patent No. 5,612,179)

130. GTG incorporates by reference each and every allegation in paragraphs 1 through 129 as though fully set forth herein.

131. ESTA has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products and/or services that infringed one or more claims of the '179 Patent in violation of 35 U.S.C. § 271(a) and/or has induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b) and/or has contributed to direct infringement of the '179 Patent by others by offering to sell, selling or providing one or more items which constitute a material part of inventions defined by claims of the '179 Patent, knowing the same to have been especially made or adapted for use in an infringement of the '179 Patent, which components are not staple articles

or commodities of commerce suitable for substantial non-infringing use in violation of 35 U.S.C. § 271(c).

132. ETSA's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

133. GTG has been damaged as a result of ESTA's infringing conduct. ESTA is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

VIII. FOURTH CLAIM FOR RELIEF
(Patent Infringement by GeneSeek – U.S. Patent No. 5,612,179)

134. GTG incorporates by reference each and every allegation in paragraphs 1 through 133 as though fully set forth herein.

135. GeneSeek has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products and/or services that infringed one or more claims of the '179 Patent in violation of 35 U.S.C. § 271(a). GeneSeek's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

136. GTG has been damaged as a result of GeneSeek's infringing conduct. GeneSeek is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

IX. FIFTH CLAIM FOR RELIEF
(Patent Infringement by GSK – U.S. Patent No. 5,612,179)

137. GTG incorporates by reference each and every allegation in paragraphs 1 through 136 as though fully set forth herein.

138. GSK has used and/or practiced the methods set forth in one or more claims of the '179 Patent, and/or has masterminded the use and/or practice of the methods set forth in one or more claims of the '179 Patent by GSK and its service providers, in violation of 35 U.S.C. § 271(a).

139. Upon information and belief, GSK was aware of and encouraged its service providers to conduct pharmacogenetic research activities in a manner that infringed the '179 Patent, and intended to cause that infringement. Alternatively, Agilent willfully blinded itself to the infringing nature of its service provider's infringing pharmacogenetic research activities. GSK has induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b).

140. GSK's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

141. GTG has been damaged as a result of GSK's infringing conduct. GSK is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

X. SIXTH CLAIM FOR RELIEF
(Patent Infringement by Hologic – U.S. Patent No. 5,612,179)

142. GTG incorporates by reference each and every allegation in paragraphs 1 through 141 as though fully set forth herein.

143. Hologic has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products and/or services that infringed one or more claims of the '179 Patent in violation of 35 U.S.C. § 271(a) and/or has induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b) and/or has contributed to direct infringement of the '179 Patent by others by offering to sell, selling or providing one or more items which constitute a material part of inventions defined by claims of the '179 Patent, knowing the same to have been especially made or adapted for use in an infringement of the '179 Patent, which components are not staple articles or commodities of commerce suitable for substantial non-infringing use in violation of 35 U.S.C. § 271(c).

144. Hologic's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

145. GTG has been damaged as a result of Hologic's infringing conduct. Hologic is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

XI. SEVENTH CLAIM FOR RELIEF
(Patent Infringement by Merial – U.S. Patent No. 5,612,179)

146. GTG incorporates by reference each and every allegation in paragraphs 1 through 145 as though fully set forth herein.

147. Merial has used and/or practiced the methods set forth in one or more claims of the '179 Patent, and/or has masterminded the use and/or practice of the methods set forth in one or more claims of the '179 Patent by Merial and its service providers, in violation of 35 U.S.C. § 271(a).

148. Upon information and belief, Merial was aware of and encourage its service providers to conduct non-coding DNA analysis methods in a manner that infringed the '179 Patent, and intended to cause that infringement. Alternatively, Merial willfully blinded itself to the infringing nature of its service providers' infringing activities. Merial has thus induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b).

149. Merial's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

150. GTG has been damaged as a result of Merial's infringing conduct. Merial is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

XII. EIGHTH CLAIM FOR RELIEF
(Patent Infringement by Pfizer – U.S. Patent No. 5,612,179)

151. GTG incorporates by reference each and every allegation in paragraphs 1 through 150 as though fully set forth herein.

152. Pfizer has used and/or practiced methods set forth in one or more claims of the '179 Patent and/or has masterminded the combined use and/or practice of the methods set forth in one or more claims of the '179 Patent by Pfizer and its service providers, in violation of 35 U.S.C. § 271(a).

153. Upon information and belief, Pfizer was aware of and encouraged laboratories and service providers to conduct analysis of non-coding DNA in a manner that infringed the '179 Patent, and intended to cause that infringement. Alternatively, Pfizer willfully blinded itself to the infringing nature of these analysis activities by laboratories and service providers. Pfizer has thus induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b).

154. Pfizer's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

155. GTG has been damaged as a result of Pfizer's infringing conduct. Pfizer is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

XIII. NINTH CLAIM FOR RELIEF
(Patent Infringement by 454 – U.S. Patent No. 5,612,179)

156. GTG incorporates by reference each and every allegation in paragraphs 1 through 155 as though fully set forth herein.

157. 454 has used and/or practiced methods set forth in one or more claims of the '179 Patent, and/or has masterminded the combined use and/or practice of the methods set forth in one or more claims of the '179 Patent by 454 and scientists acting at 454's direction, in violation of 35 U.S.C. § 271(a).

158. 454 has manufactured, made, had made, used, practiced, imported, provided, supplied, distributed, sold, and/or offered for sale products, the use of which by 454's customers infringed one or more claims of the '179 Patent. Upon information and belief, 454 was aware of and encouraged its customers to use 454 products in a manner that infringed the '179 Patent, and intended to cause that infringement. Alternatively, 454 willfully blinded itself to the infringing nature of its customer's use of 454 products to infringe the '179 Patent. 454 has thus induced direct infringement of the '179 Patent by others by actively instructing, assisting and/or encouraging others to practice one or more of the inventions claimed in the '179 Patent in violation of 35 U.S.C. § 271(b).

159. 454's actions in infringing the '179 Patent have been, and are, willful, deliberate and/or in conscious disregard of GTG's rights, making this an exceptional case within the meaning of 35 U.S.C. § 285 and entitling GTG to the award of its attorneys' fees.

160. GTG has been damaged as a result of 454's infringing conduct. 454 is thus liable to GTG in an amount that adequately compensates GTG for such infringement which cannot be

less than a reasonable royalty, together with interest and costs as fixed by this Court under 35 U.S.C. § 284.

XIV. JURY DEMAND

GTG hereby requests a trial by jury pursuant to Rule 38 of the Federal Rules of Civil Procedure.

XV. PRAYER FOR RELIEF

GTG requests that the Court find in its favor and against Defendants, and that the Court grant GTG the following relief:

A. Judgment that one or more claims of the '179 Patent has been directly infringed, either literally, and/or under the doctrine of equivalents, by one or more Defendants and/or by others to whose infringement Defendants have contributed and/or by others whose infringement has been induced by Defendants;

B. Judgment that Defendants account for and pay to GTG all damages to and costs incurred by GTG because of Defendants' infringing activities and other conduct complained of herein in an amount not less than a reasonable royalty;

C. That such damages be trebled where allowed by law for a Defendants' willful infringement;

D. That GTG be granted pre-judgment and post-judgment interest on the damages caused to it by reason of Defendants' infringing activities and other conduct complained of herein;

E. That this Court declare this an exceptional case and award GTG its reasonable attorney's fees and costs in accordance with 35 U.S.C. § 285; and

F. That GTG be granted such other and further relief as the court may deem just and proper under the circumstances.

Respectfully submitted,

Dated: September 16, 2011

By: s/ Benjamin B. Lieb

Robert R. Brunelli
rbrunelli@sheridanross.com
Todd P. Blakely
tblakely@sheridanross.com
Benjamin B. Lieb
blieb@sheridanross.com
Hiwot M. Covell
hcovell@sheridanross.com
SHERIDAN ROSS P.C.
1560 Broadway, Suite 1200
Denver, Colorado 80202-5141
(303) 863-9700
(303) 863-0223 (facsimile)
litigation@sheridanross.com

ATTORNEYS FOR PLAINTIFF

US005612179A

United States Patent [19]

Simons

[11] Patent Number: 5,612,179

[45] Date of Patent: *Mar. 18, 1997

[54] **INTRON SEQUENCE ANALYSIS METHOD FOR DETECTION OF ADJACENT AND REMOTE LOCUS ALLELES AS HAPLOTYPES**

[75] Inventor: **Malcolm J. Simons**, Fryerstown, New Zealand

[73] Assignee: **GeneType A.G.**, Zug, Switzerland

[*] Notice: The portion of the term of this patent subsequent to Mar. 9, 2010, has been disclaimed.

[21] Appl. No.: **949,652**

[22] Filed: **Sep. 23, 1992**

Related U.S. Application Data

[63] Continuation of Ser. No. 551,239, Jul. 11, 1990, Pat. No. 5,192,659, which is a continuation-in-part of Ser. No. 465,863, Jan. 16, 1990, abandoned, which is a continuation-in-part of Ser. No. 405,499, Sep. 11, 1989, abandoned, which is a continuation-in-part of Ser. No. 398,217, Aug. 25, 1989, abandoned.

[51] **Int. Cl.**[6] **C12Q 1/68**; C12P 19/34; C07H 21/04; C12N 15/00

[52] **U.S. Cl.** **435/6**; 435/91.1; 435/91.2; 536/23.1; 536/24.3; 536/24.31; 536/24.33; 935/77; 935/78

[58] **Field of Search** 435/91, 6, 91.1, 435/91.5, 91.2; 935/77, 78; 536/24.31, 24.32, 23.1, 23.5, 25.3

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,582,788	4/1986	Erlich	435/6
4,683,194	7/1987	Saiki	435/6
4,683,195	7/1987	Mullis et al.	435/6
5,192,659	3/1993	Simons	435/6

FOREIGN PATENT DOCUMENTS

873051007	6/1987	European Pat. Off. .
873094825	10/1987	European Pat. Off. .
893104240	10/1988	European Pat. Off. .
US89/02169	5/1989	WIPO .
US89/02731	6/1989	WIPO .

OTHER PUBLICATIONS

Waikan et al. P.N.A.S, U.S.A., vol. 75, No. 11, pp. 5631–5635, Nov. 1978.

Allen et al., *Biotechniques*, vol. 7, No. 7, pp. 736–744 (1989).

DiLella, et al., "Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria," *Nature*, vol. 322, No. 28, Aug. 1986, pp. 799–803.

DiLella, et al., "Screening for Phenylketonuria Mutations by DNA Amplification with the Polymerase Chain Reaction," *The Lancet*, Mar. 5, 1988, pp.. 497–499.

Clark, "Interference of Haplotypes from PCR-amplified Samples of Diploid Populations", *Mol. Biol. Evol.*, vol. 7, No. 2, 1990, pp. 111–122.

Dicker, et al., "Sequence Analysis of a Human Gene Responsible for Drug Resistance: A Rapid Method for Manual and Automated Direct Sequencing of Products Generated by the Polymerase Chain Reaction," *BioTechniques*, vol. 7, No. 8, 1989, pp. 830–837.

Kogan et al., "An Improved Method for Prenatal Diagnosis of Genetic Diseases by Analysis of Amplifed DNA Sequences—Application to Hemophilia A," The New England Journal of Medicine, vol. 317, No. 16, Oct. 15, 1987, pp. 985–990.

Mardis et al., "Automated Methods for Single–Stranded DNA Isolation and Dideoxynucleotide DNA Sequencing Reactions on a Robotic Workstation," *BioTechniques*, vol. 7, No. 8, 1989, pp. 840–850.

Marx, "Disecting the Complex Diseases," *Science*, vol. 247, Mar. 30, 1990, pp. 1540–1542.

Riordan et al., "Identification of the Cystic Fibrosis Gene: Cloning and Characterization of Complementary DNA,", *Science*, vol. 245, 1989, pp. 1066–1073.

Rommens et al., "Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping,", *Science*, vol. 245, 1989, pp. 1059–1065.

Stephens et al., "Theoretical Underpinning of the Single–Molecule–Dilution (SMD) Method of Direct Haplotype Resolution," *Am. J. Hum. Genet*, vol. 46, 1990, pp. 1149–1155.

Graham et al., "Application of Molecular Genetics to Prenatal Diagnosis and Carrier Detection in the Hemophilias: Some Limitaitons," *Blood*, vol. 66, No. 4, Oct. 1985, pp. 759–764.

Saiki et al., "Enzymatic Amplification of β–Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia,", *Science*, vol. 230, 1985, pp. 1350–1354.

Funke et al., "Detection of a New Msp I Restriction Fragment Length Polymorphism in the Apolipoprotein A–I Gene," *J. Clin. Chem. Clin. Biochem.*, vol. 25, No. 3, 1987, pp. 131–134.

(List continued on next page.)

Primary Examiner—W. Gary Jones
Assistant Examiner—Bradley L. Sisson
Attorney, Agent, or Firm—Skjerven, Morrill, MacPherson, Franklin & Friel; Laura Terlizzi

[57] **ABSTRACT**

The present invention provides a method for detection of at least one allele of a genetic locus and can be used to provide direct determination of the haplotype. The method comprises amplifying genomic DNA with a primer pair that spans an intron sequence and defines a DNA sequence in genetic linkage with an allele to be detected. The primer-defined DNA sequence contains a sufficient number of intron sequence nucleotides to characterize the allele. Genomic DNA is amplified to produce an amplified DNA sequence characteristic of the allele. The amplified DNA sequence is analyzed to detect the presence of a genetic variation in the amplified DNA sequence such as a change in the length of the sequence, gain or loss of a restriction site or substitution of a nucleotide. The variation is characteristic of the allele to be detected and can be used to detect remote alleles. Kits comprising one or more of the reagents used in the method are also described.

36 Claims, No Drawings

OTHER PUBLICATIONS

Erlich et al., "Reliability of the HLA–DQα PCR–Based Olignucleotide Typing System," *Journal of Forensic Sciences*, Sep. 1990, pp. 1017–1018.

Boehnke et al., "Fine–Structure Genetic Mapping of Human Chromosomes Using the Polymerase Chain Reaction on Single Sperm: Experimental Design Considerations," *The American Journal of Human Genetics*, Volumr 45, No. 1, Jul. 1989, pp. 21–32.

Geisel et al., "A New APA LI Restriction Fragment Length Polymrophism in the Low Density Lipoprotein Receptor Gene," *Journal of Clinical Chemistry and Clinical Biochemistry*, vol. 26, No. 7, 1988, pp. 429–434.

Kerem et al., "Identification of the Cystic Fibrosis Gene: Genetic Analysis,", *Science*, vol. 245, 1989, pp. 1073–1080.

Deng, "A sensitive non–radioactive PCR–RFLP analysis for detecting point mutations at 12th condon oncogene c–Ha–ras in DNAs of gastic cancer," *Nucleic Acids Research*, vol. 16, No. 13, 1988, p. 6231.

Riess, et al., "Hypervariability of intronic simple $(gt)_n(ga)_m$ repeats in HLA–DRB genes," *Immunogenetics*, vol. 32, No. 2, Aug. 1990, pp. 110–116.

Abraham, et al., "Sequence differences between HLA–B and TNF distinguish different MHC ancestral haplotypes," *Antigens*, vol 39, No. 3, Mar. 1992, pp. 117–121.

Simons et al., pp. 952–958 In: *Immunology of HLA*, vol. 1:Springer–Verlag, New York (1989).

Simons, et al., pp. 959–1023 In: *Immunology of HLA*, vol. 1: Springer–Verlag, New York (1989).

Molecular Biology of the Gene, 4th Edition, California, 1987, Chapter 20: "Conserved Sequences at Exon–Intron Boundaries in Pre–mRNA," pp. 639–640.

Molecular Biology of The Cell, 2nd Edition, New York, 1989, Chapter 9: "RNA Synthesis and RNA Processing" pp. 531–535 and Chapter 10: Control of Gene Expression pp. 602–604.

Simon and Erlich, pp. 952–958 In: *Immunology of HLA*, vol. 1: Springer–Verlag, New York (1989).

Padgett et al. *Amer. Rev. Biochem.* 1986. 55:1119–150.

INTRON SEQUENCE ANALYSIS METHOD FOR DETECTION OF ADJACENT AND REMOTE LOCUS ALLELES AS HAPLOTYPES

CROSS REFERENCE TO RELATED APPLICATION

This application is a continuation of application Ser. No. 07/551,239, filed Jul. 11, 1990. This application is a continuation of U.S. Pat. No. 5,192,659, issued Mar. 9, 1993, (U.S. Ser. No. 07/551,239, filed Jul. 11, 1990), which is a continuation-in-part of U.S. Ser. No. 07/465,863, filed Jan. 16, 1990, which is a continuation in part of U.S. Ser. No. 07/405,499, filed Sep. 11, 1989 (now abandoned), which is a continuation in part of U.S. application Ser. No. 07/398,217, filed Aug. 25, 1989 (now abandoned).

FIELD OF THE INVENTION

The present invention relates to a method for detection of alleles and haplotypes and reagents therefor.

BACKGROUND OF THE INVENTION

Due in part to a number of new analytical techniques, there has been a significant increase in knowledge about genetic information, particularly in humans. Allelic variants of genetic loci have been correlated to malignant and non-malignant monogenic and multigenic diseases. For example, monogenic diseases for which the defective gene has been identified include DuChenne muscular dystrophy, sickle-cell anemia, Lesch Nyhan syndrome, hemophilia, beta-thalassemia, cystic fibrosis, polycystic kidney disease, ADA deficiency, α-1-antitrypsin deficiency, Wilm's tumor and retinoblastoma. Other diseases which are believed to be monogenic for which the gene has not been identified include fragile X mental retardation and Huntington's chorea.

Genes associated with multigenic diseases such as diabetes, colon cancer and premature coronary atherosclerosis have also been identified.

In addition to identifying individuals at risk for or carriers of genetic diseases, detection of allelic variants of a genetic locus has been used for organ transplantation, forensics, disputed paternity and a variety of other purposes in humans. In commercially important plants and animals, genes have not only been analyzed but genetically engineered and transmitted into other organisms.

A number of techniques have been employed to detect allelic variants of genetic loci including analysis of restriction fragment length polymorphic (RFLP) patterns, use of oligonucleotide probes, and DNA amplification methods. One of the most complicated groups of allelic variants, the major histocompatibility complex (MHC), has been extensively studied. The problems encountered in attempting to determine the HLA type of an individual are exemplary of problems encountered in characterizing other genetic loci.

The major histocompatibility complex is a cluster of genes that occupy a region on the short arm of chromosome 6. This complex, denoted the human leukocyte antigen (HLA) complex, includes at least 50 loci. For the purposes of HLA tissue typing, two main classes of loci are recognized. The Class I loci encode transplantation antigens and are designated A, B and C. The Class II loci (DRA, DRB, DQA1, DQB, DPA and DPB) encode products that control immune responsiveness. Of the Class II loci, all the loci are

polymorphic with the exception of the DRA locus. That is, the DRA antigen polypeptide sequence is invariant.

HLA determinations are used in paternity determinations, transplant compatibility testing, forensics, blood component therapy, anthropological studies, and in disease association correlations to diagnose disease or predict disease susceptibility. Due power of HLA to distinguish individuals and the need to match HLA type for transplantation, analytical methods to unambiguously characterize the alleles of the genetic loci associated with the complex have been sought. At present, DNA typing using RFLP and oligonucleotide probes has been used to type Class II locus alleles. Alleles of Class I loci and Class II DR and DQ loci are typically determined by serological methods. The alleles of the Class II DP locus are determined by primed lymphocyte typing (PLT).

Each of the HLA analysis methods has drawbacks. Serological methods require standard sera that are not widely available and must be continuously replenished. Additionally, serotyping is based on the reaction of the HLA gene products in the sample with the antibodies in the reagent sera. The antibodies recognize the expression products of the HLA genes on the surface of nucleated cells. The determination of fetal HLA type by serological methods may be difficult due to lack of maturation of expression of the antigens in fetal blood cells.

Oligonucleotide probe typing can be performed in two days and has been further improved by the recent use of polymerase chain reaction (PCR) amplification. PCR-based oligoprobe typing has been performed on Class II loci. Primed lymphocyte typing requires 5 to 10 days to complete and involves cell culture with its difficulties and inherent variability.

RFLP analysis is time consuming, requiring about 5 to 7 days to complete. Analysis of the fragment patterns is complex. Additionally, the technique requires the use of labelled probes. The most commonly used label, ^{32}P, presents well known drawbacks associated with the use of radionuclides.

A fast, reliable method of genetic locus analysis is highly desirable.

DESCRIPTION OF THE PRIOR ART

U.S. Pat. No. 4,683,195 (to Mullis et al, issued Jul. 28, 1987) describes a process for amplifying, detecting and/or cloning nucleic acid sequences. The method involves treating separate complementary strands of DNA with two oligonucleotide primers, extending the primers to form complementary extension products that act as templates for synthesizing the desired nucleic acid sequence and detecting the amplified sequence. The method is commonly referred to as the polymerase chain reaction sequence amplification method or PCR. Variations of the method are described in U.S. Pat. No. 4,683,194 (to Saiki et al, issued Jul. 28, 1987). The polymerase chain reaction sequence amplification method is also described by Saiki et al, *Science,* 230:1350–1354 (1985) and Scharf et al, *Science,* 324:163–166 (1986).

U.S. Pat. No. 4,582,788 (to Erlich, issued Apr. 15, 1986) describes an HLA typing method based on restriction length polymorphism (RFLP) and cDNA probes used therewith. The method is carried out by digesting an individual's HLA DNA with a restriction endonuclease that produces a polymorphic digestion pattern, subjecting the digest to genomic blotting using a labelled cDNA probe that is complementary

to an HLA DNA sequence involved in the polymorphism, and comparing the resulting genomic blotting pattern with a standard. Locus-specific probes for Class II loci (DQ) are also described.

Kogan et al, *New Engl. J. Med,* 317:985–990 (1987) describes an improved PCR sequence amplification method that uses a heat-stable polymerase (Taq polymerase) and high temperature amplification. The stringent conditions used in the method provide sufficient fidelity of replication to permit analysis of the amplified DNA by determining DNA sequence lengths by visual inspection of an ethidium bromide-stained gel. The method was used to analyze DNA associated with hemophilia A in which additional tandem repeats of a DNA sequence are associated with the disease and the amplified sequences were significantly longer than sequences that are not associated with the disease.

Simons and Erlich, pp 952–958 In: *Immunology of HLA* Vol. 1: Springer-Verlag, New York (1989) summarized RFLP-sequence interrelations at the DPA and DPB loci. RFLP fragment patterns analyzed with probes by Southern blotting provided distinctive patterns for DPw1-5 alleles and the corresponding DPB1 allele sequences, characterized two subtypic patterns for DPw2 and DPw4, and identified new DPw alleles.

Simons et al, pp 959–1023 In: *Immunology of HLA* Vol. 1: Springer-Verlag, New York (1989) summarized restriction length polymorphisms of HLA sequences for class II loci as determined by the 10th International Workshop Southern Blot Analysis. Southern blot analysis was shown to be suitable for typing of the major classes of HLA loci.

A series of three articles [Rommens et al, *Science* 245:1059–1065 (1989), Riordan et al, *Science* 245:1066–1072 (1989) and Kerem et al, *Science* 245:1073–1079 (1989) report a new gene analysis method called "jumping" used to identify the location of the CF gene, the sequence of the CF gene, and the defect in the gene and its percentage in the disease population, respectively.

DiLelia et al, *The Lancet* i:497–499 (1988) describes a screening method for detecting the two major alleles responsible for phenylketonuria in caucasians of Northern European descent. The mutations, located at about the center of exon 12 and at the exon 12 junction with intervening sequence 12 are detected by PCR amplification of a 245 bp region of exon 12 and flanking intervening sequences. The amplified sequence encompasses both mutations and is analyzed using probes specific for each of the alleles (without prior electrophoretic separation).

Dicker et al, *BioTechniques* 7:830–837 (1989) and Mardis et al, *BioTechniques* 7:840–850 (1989) report on automated techniques for sequencing of DNA sequences, particularly PCR-generated sequences.

Each of the above-described references is incorporated herein by reference in its entirety.

SUMMARY OF THE INVENTION

The present invention provides a method for detection of at least one allele of a genetic locus and can be used to provide direct determination of the haplotype. The method comprises amplifying genomic DNA with a primer pair that spans an intron sequence and defines a DNA sequence in genetic linkage with an allele to be detected. The primer-defined DNA sequence contains a sufficient number of intron sequence nucleotides to characterize the allele. Genomic DNA is amplified to produce an amplified DNA sequence characteristic of the allele. The amplified DNA sequence is analyzed to detect the presence of a genetic variation in the amplified DNA sequence such as a change in the length of the sequence, gain or loss of a restriction site or substitution of a nucleotide. The variation is characteristic of the allele to be detected.

The present invention is based on the finding that intron sequences contain genetic variations that are characteristic of adjacent and remote alleles on the same chromosome. In particular, DNA sequences that include a sufficient number of intron sequence nucleotides can be used for direct determination of haplotype.

The method can be used to detect alleles of genetic loci for any eukaryotic organism. Of particular interest are loci associated with malignant and nonmalignant monogenic and multigenic diseases, and identification of individual organisms or species in both plants and animals. In a preferred embodiment, the method is used to determine HLA allele type and haplotype.

Kits comprising one or more of the reagents used in the method are also described.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a method for detection of alleles and haplotypes through analysis of intron sequence variation. The present invention is based on the discovery that amplification of intron sequences that exhibit linkage disequilibrium with adjacent and remote loci can be used to detect alleles of those loci. The present method reads haplotypes as the direct output of the intron typing analysis when a single, individual organism is tested. The method is particularly useful in humans but is generally applicable to all eukaryotes, and is preferably used to analyze plant and animal species.

The method comprises amplifying genomic DNA with a primer pair that spans an intron sequence and defines a DNA sequence in genetic linkage with an allele to be detected. Primer sites are located in conserved regions in the introns or exons bordering the intron sequence to be amplified. The primer-defined DNA sequence contains a sufficient number of intron sequence nucleotides to characterize the allele. The amplified DNA sequence is analyzed to detect the presence of a genetic variation such as a change in the length of the sequence, gain or loss of a restriction site or substitution of a nucleotide.

The intron sequences provide genetic variations that, in addition to those found in exon sequences, further distinguish sample DNA, providing additional information about the individual organism. This information is particularly valuable for identification of individuals such as in paternity determinations and in forensic applications. The information is also valuable in any other application where heterozygotes (two different alleles) are to be distinguished from homozygotes (two copies of one allele).

More specifically, the present invention provides information regarding intron variation. Using the methods and reagents of this invention, two types of intron variation associated with genetic loci have been found. The first is allele-associated intron variation. That is, the intron variation pattern associates with the allele type at an adjacent locus. The second type of variation is associated with remote alleles (haplotypes). That is, the variation is present in individual organisms with the same genotype at the primary locus. Differences may occur between sequences of the same

adjacent and remote locus types. However, individual-limited variation is uncommon.

Furthermore, an amplified DNA sequence that contains sufficient intron sequences will vary depending on the allele present in the sample. That is, the introns contain genetic variations (e.g. length polymorphisms due to insertions and/or deletions and changes in the number or location of restriction sites) which are associated with the particular allele of the locus and with the alleles at remote loci.

The reagents used in carrying out the methods of this invention are also described. The reagents can be provided in kit form comprising one or more of the reagents used in the method.

Definitions

The term "allele", as used herein, means a genetic variation associated with a coding region; that is, an alternative form of the gene.

The term "linkage", as used herein, refers to the degree to which regions of genomic DNA are inherited together. Regions on different chromosomes do not exhibit linkage and are inherited together 50% of the time. Adjacent genes that are always inherited together exhibit 100% linkage.

The term "linkage disequilibrium", as used herein, refers to the co-occurrence of two alleles at linked loci such that the frequency of the co-occurrence of the alleles is greater than would be expected from the separate frequencies of occurrence of each allele. Alleles that co-occur with frequencies expected from their separate frequencies are said to be in "linkage equilibrium".

As used herein, "haplotype" is a region of genomic DNA on a chromosome which is bounded by recombination sites such that genetic loci within a haplotypic region are usually inherited as a unit. However, occasionally, genetic rearrangements may occur within a haplotype. Thus, the term haplotype is an operational term that refers to the occurrence on a chromosome of linked loci.

As used herein, the term "intron" refers to untranslated DNA sequences between exons, together with 5' and 3' untranslated regions associated with a genetic locus. In addition, the term is used to refer to the spacing sequences between genetic loci (intergenic spacing sequences) which are not associated with a coding region and are colloquially referred to as "junk". While the art traditionally uses the term "intron" to refer only to untranslated sequences between exons, this expanded definition was necessitated by the lack of any art recognized term which encompasses all non-exon sequences.

As used herein, an "intervening sequence" is an intron which is located between two exons within a gene. The term does not encompass upstream and downstream noncoding sequences associated with the genetic locus.

As used herein, the term "amplified DNA sequence" refers to DNA sequences which are copies of a portion of a DNA sequence and its complementary sequence, which copies correspond in nucleotide sequence to the original DNA sequence and its complementary sequence.

The term "complement", as used herein, refers to a DNA sequence that is complementary to a specified DNA sequence.

The term "primer site", as used herein, refers to the area of the target DNA to which a primer hybridizes.

The term "primer pair", as used herein, means a set of primers including a 5' upstream primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3' downstream primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

The term "exon-limited primers", as used herein, means a primer pair having primers located within or just outside of an exon in a conserved portion of the intron, which primers amplify a DNA sequence which includes an exon or a portion thereof and not more than a small, para-exon region of the adjacent intron(s).

The term "intron-spanning primers", as used herein, means a primer pair that amplifies at least a portion of one intron, which amplified intron region includes sequences which are not conserved. The intron-spanning primers can be located in conserved regions of the introns or in adjacent, upstream and/or downstream exon sequences.

The term "genetic locus", as used herein, means the region of the genomic DNA that includes the gene that encodes a protein including any upstream or downstream transcribed noncoding regions and associated regulatory regions. Therefore, an HLA locus is the region of the genomic DNA that includes the gene that encodes an HLA gene product.

As used herein, the term "adjacent locus" refers to either (1) the locus in which a DNA sequence is located or (2) the nearest upstream or downstream genetic locus for intron DNA sequences not associated with a genetic locus.

As used herein, the term "remote locus" refers to either (1) a locus which is upstream or downstream from the locus in which a DNA sequence is located or (2) for intron sequences not associated with a genetic locus, a locus which is upstream or downstream from the nearest upstream or downstream genetic locus to the intron sequence.

The term "locus-specific primer", as used herein, means a primer that specifically hybridizes with a portion of the stated gene locus or its complementary strand, at least for one allele of the locus, and does not hybridize with other DNA sequences under the conditions used in the amplification method.

As used herein, the terms "endonuclease" and "restriction endonuclease" refer to an enzyme that cuts double-stranded DNA having a particular nucleotide sequence. The specificities of numerous endonucleases are well known and can be found in a variety of publications, e.g. *Molecular Cloning: A Laboratory Manual* by Maniatis et al, Cold Spring Harbor Laboratory 1982. That manual is incorporated herein by reference in its entirety.

The term "restriction fragment length polymorphism" (or RFLP), as used herein, refers to differences in DNA nucleotide sequences that produce fragments of different lengths when cleaved by a restriction endonuclease.

The term "primer-defined length polymorphisms" (or PDLP), as used herein, refers to differences in the lengths of amplified DNA sequences due to insertions or deletions in the intron region of the locus included in the amplified DNA sequence.

The term "HLA DNA", as used herein, means DNA that includes the genes that encode HLA antigens. HLA DNA is found in all nucleated human cells.

Primers

The method of this invention is based on amplification of selected intron regions of genomic DNA. The methodology is facilitated by the use of primers that selectively amplify DNA associated with one or more alleles of a genetic locus of interest and not with other genetic loci.

A locus-specific primer pair contains a 5' upstream primer that defines the 5' end of the amplified sequence by hybridizing with the 5' end of the target sequence to be amplified and a 3' downstream primer that defines the 3' end of the amplified sequence by hybridizing with the complement of the 3' end of the DNA sequence to be amplified. The primers in the primer pair do not hybridize with DNA of other genetic loci under the conditions used in the present invention.

For each primer of the locus-specific primer pair, the primer hybridizes to at least one allele of the DNA locus to be amplified or to its complement. A primer pair can be prepared for each allele of a selected locus, which primer pair amplifies only DNA for the selected locus. In this way combinations of primer pairs can be used to amplify genomic DNA of a particular locus, irrespective of which allele is present in a sample. Preferably, the primer pair amplifies DNA of at least two, more preferably more than two, alleles of a locus. In a most preferred embodiment, the primer sites are conserved, and thus amplify all haplotypes. However, primer pairs or combinations thereof that specifically bind with the most common alleles present in a particular population group are also contemplated.

The amplified DNA sequence that is defined by the primers contains a sufficient number of intron sequence nucleotides to distinguish between at least two alleles of an adjacent locus, and preferably, to identify the allele of the locus which is present in the sample. For some purposes, the sequence can also be selected to contain sufficient genetic variations to distinguish between individual organisms with the same allele or to distinguish between haplotypes.

Length of sequence

The length of the amplified sequence which is required to include sufficient genetic variability to enable discrimination between all alleles of a locus bears a direct relation to the extent of the polymorphism of the locus (the number of alleles). That is, as the number of alleles of the tested locus increases, the size of an amplified sequence which contains sufficient genetic variations to identify each allele increases. For a particular population group, one or more of the recognized alleles for any given locus may be absent from that group and need not be considered in determining a sequence which includes sufficient variability for that group. Conveniently, however, the primer pairs are selected to amplify a DNA sequence which is sufficient to distinguish between all recognized alleles of the tested locus. The same considerations apply when a haplotype is determined.

For example, the least polymorphic HLA locus is DPA which currently has four recognized alleles. For that locus, a primer pair which amplifies only a portion of the variable exon encoding the allelic variation contains sufficient genetic variability to distinguish between the alleles when the primer sites are located in an appropriate region of the variable exon. Exon-limited primers can be used to produce an amplified sequence that includes as few as about 200 nucleotides (nt). However, as the number of alleles of the locus increases, the number of genetic variations in the sequence must increase to distinguish all alleles. Addition of invariant exon sequences provides no additional genetic variation. When about eight or more alleles are to be distinguished, as for the DQA1 locus and more variable loci, amplified sequences should extend into at least one intron in the locus, preferably an intron adjacent to the variable exon.

Additionally, where alleles of the locus exist which differ by a single basepair in the variable exon, intron sequences are included in amplified sequences to provide sufficient variability to distinguish alleles. For example, for the DQA1 locus (with eight currently recognized alleles) and the DPB locus (with 24 alleles), the DQA1.1/1.2 (now referred to as DQA1 0101/0102) and DPB2.1/4.2 (now referred to as DPB0201/0402) alleles differ by a single basepair. To distinguish those alleles, amplified sequences which include an intron sequence region are required. About 300 to 500 nucleotides is sufficient, depending on the location of the sequence. That is, 300 to 500 nucleotides comprised primarily of intron sequence nucleotides sufficiently close to the variable exon are sufficient.

For loci with-more extensive polymorphisms (such as DQB with 14 currently recognized alleles, DPB with 24 currently recognized alleles, DRB with 34 currently recognized alleles and for each of the Class I loci), the amplified sequences need to be larger to provide sufficient variability to distinguish between all the alleles. An amplified sequence that includes at least about 0.5 kilobases (Kb), preferably at least about 1.0 Kb, more preferably at least about 1.5 Kb generally provides a sufficient number of restriction sites for loci with extensive polymorphisms. The amplified sequences used to characterize highly polymorphic loci are generally between about 800 to about 2,000 nucleotides (nt), preferably between about 1000 to about 1800 nucleotides in length.

When haplotype information regarding remote alleles is desired, the sequences are generally between about 1,000 to about 2,000 nt in length. Longer sequences are required when the amplified sequence encompasses highly conserved regions such as exons or highly conserved intron regions, e.g., promoters, operators and other DNA regulatory regions. Longer amplified sequences (including more intron nucleotide sequences) are also required as the distance between the amplified sequences and the allele to be detected increases.

Highly conserved regions included in the amplified DNA sequence, such as exon sequences or highly conserved intron sequences (e.g. promoters, enhancers, or other regulatory regions) may provide little or no genetic variation. Therefore, such regions do not contribute, or contribute only minimally, to the genetic variations present in the amplified DNA sequence. When such regions are included in the amplified DNA sequence, additional nucleotides may be required to encompass sufficient genetic variations to distinguish alleles, in comparison to an amplified DNA sequence of the same length including only intron sequences.

Location of the amplified DNA sequence

The amplified DNA sequence is located in a region of genomic DNA that contains genetic variation which is in genetic linkage with the allele to be detected. Preferably, the sequence is located in an intron sequence adjacent to an exon of the genetic locus. More preferably, the amplified sequence includes an intervening sequence adjacent to an exon that encodes the allelic variability associated with the locus (a variable exon). The sequence preferably includes at least a portion of one of the introns adjacent to a variable exon and can include a portion of the variable exon. When additional sequence information is required, the amplified DNA sequence preferably encompasses a variable exon and all or a portion of both adjacent intron sequences.

Alternatively, the amplified sequence can be in an intron which does not border an exon of the genetic locus. Such introns are located in the downstream or upstream gene

flanking regions or even in an intervening sequence in another genetic locus which is in linkage disequilibrium with the allele to be detected.

For some genetic loci, genomic DNA sequences may not be available. When only cDNA sequences are available and intron locations within the sequence are not identified, primers are selected at intervals of about 200 nt and used to amplify genomic DNA. If the amplified sequence contains about 200 nt, the location of the first primer is moved about 200 nt to one side of the second primer location and the amplification is repeated until either (1) an amplified DNA sequence that is larger than expected is produced or (2) no amplified DNA sequence is produced. In either case, the location of an intron sequence has been determined. The same methodology can be used when only the sequence of a marker site that is highly linked to the genetic locus is available, as is the case for many genes associated with inherited diseases.

When the amplified DNA sequence does not include all or a portion of an intron adjacent to the variable exon(s), the sequence must also satisfy a second requirement. The amplified sequence must be sufficiently close to the variable exon(s) to exclude recombination and loss of linkage disequilibrium between the amplified sequence and the variable exon(s). This requirement is satisfied if the regions of the genomic DNA are within about 5 Kb, preferably within about 4 Kb, most preferably within 2 Kb of the variable exon(s). The amplified sequence can be outside of the genetic locus but is preferably within the genetic locus.

Preferably, for each primer pair, the amplified DNA sequence defined by the primers includes at least 200 nucleotides, and more preferably at least 400 nucleotides, of an intervening sequence adjacent to the variable exon(s). Although the variable exon usually provides fewer variations in a given number of nucleotides than an adjacent intervening sequence, each of those variations provides allele-relevant information. Therefore, inclusion of the variable exon provides an advantage.

Since PCR methodology can be used to amplify sequences of several Kb, the primers can be located so that additional exons or intervening sequences are included in the amplified sequence. Of course, the increased size of the amplified DNA sequence increases the chance of replication error, so addition of invariant regions provides some disadvantages. However, those disadvantages are not as likely to affect an analysis based on the length of the sequence or the RFLP fragment patterns as one based on sequencing the amplification product. For particular alleles, especially those with highly similar exon sequences, amplified sequences of greater than about 1 or 1.5 Kb may be necessary to discriminate between all alleles of a particular locus.

The ends of the amplified DNA sequence are defined by the primer pair used in the amplification. Each primer sequence must correspond to a conserved region of the genomic DNA sequence. Therefore, the location of the amplified sequence will, to some extent, be dictated by the need to locate the primers in conserved regions. When sufficient intron sequence information to determine conserved intron regions is not available, the primers can be located in conserved portions of the exons and used to amplify intron sequences between those exons.

When appropriately-located, conserved sequences are not unique to the genetic locus, a second primer located within the amplified sequence produced by the first primer pair can be used to provide an amplified DNA sequence specific for the genetic locus. At least one of the primers of the second primer pair is located in a conserved region of the amplified DNA sequence defined by the first primer pair. The second primer pair is used following amplification with the first primer pair to amplify a portion of the amplified DNA sequence produced by the first primer pair.

There are three major types of genetic variations that can be detected and used to identify an allele. Those variations, in order of ease of detection, are (1) a change in the length of the sequence, (2) a change in the presence or location of at least one restriction site and (3) the substitution of one or a few nucleotides that does not result in a change in a restriction site. Other variations within the amplified DNA sequence are also detectable.

There are three types of techniques which can be used to detect the variations. The first is sequencing the amplified DNA sequence. Sequencing is the most time consuming and also the most revealing analytical method, since it detects any type of genetic variation in the amplified sequence. The second analytical method uses allele-specific oligonucleotide or sequence-specific oligonucleotides probes (ASO or SSO probes). Probes can detect single nucleotide changes which result in any of the types of genetic variations, so long as the exact sequence of the variable site is known. A third type of analytical method detects sequences of different lengths (e.g., due to an insertion or deletion or a change in the location of a restriction site) and/or different numbers of sequences (due to either gain or loss of restriction sites). A preferred detection method is by gel or capillary electrophoresis. To detect changes in the lengths of fragments or the number of fragments due to changes in restriction sites, the amplified sequence must be digested with an appropriate restriction endonuclease prior to analysis of fragment length patterns.

The first genetic variation is a difference in the length of the primer-defined amplified DNA sequence, referred to herein as a primer-defined length polymorphism (PDLP), which difference in length distinguishes between at least two alleles of the genetic locus. The PDLPs result from insertions or deletions of large stretches (in comparison to the total length of the amplified DNA sequence) of DNA in the portion of the intron sequence defined by the primer pair. To detect PDLPs, the amplified DNA sequence is located in a region containing insertions or deletions of a size that is detectable by the chosen method. The amplified DNA sequence should have a length which provides optimal resolution of length differences. For electrophoresis, DNA sequences of about 300 to 500 bases in length provide optimal resolution of length differences. Nucleotide sequences which differ in length by as few as 3 nt, preferably 25 to 50 nt, can be distinguished. However, sequences as long as 800 to 2,000 nt which differ by at least about 50 nt are also readily distinguishable. Gel electrophoresis and capillary electrophoresis have similar limits of resolution. Preferably the length differences between amplified DNA sequences will be at least 10, more preferably 20, most preferably 50 or more, nt between the alleles. Preferably, the amplified DNA sequence is between 300 to 1,000 nt and encompasses length differences of at least 3, preferably 10 or more nt.

Preferably, the amplified sequence is located in an area which provides PDLP sequences that distinguish most or all of the alleles of a locus. An example of PDLP-based identification of five of the eight DQA1 alleles is described in detail in the examples.

When the variation to be detected is a change in a restriction site, the amplified DNA sequence necessarily

contains at least one restriction site which (1) is present in one allele and not in another, (2) is apparently located in a different position in the sequence of at least two alleles, or (3) combinations thereof. The amplified sequence will preferably be located such that restriction endonuclease cleavage produces fragments of detectably different lengths, rather than two or more fragments of approximately the same length.

For allelic differences detected by ASO or SSO probes, the amplified DNA sequence includes a region of from about 200 to about 400 nt which is present in one or more alleles and not present in one or more other alleles. In a most preferred embodiment, the sequence contains a region detectable by a probe that is present in only one allele of the genetic locus. However, combinations of probes which react with some alleles and not others can be used to characterize the alleles.

For the method described herein, it is contemplated that use of more than one amplified DNA sequence and/or use of more than one analytical method per amplified DNA sequence may be required for highly polymorphic loci, particularly for loci where alleles differ by single nucleotide substitutions that are not unique to the allele or when information regarding remote alleles (haplotypes) is desired. More particularly, it may be necessary to combine a PDLP analysis with an RFLP analysis, to use two or more amplified DNA sequences located in different positions or to digest a single amplified DNA sequence with a plurality of endonucleases to distinguish all the alleles of some loci. These combinations are intended to be included within the scope of this invention.

For example, the analysis of the haplotypes of DQA1 locus described in the examples uses PDLPs and RFLP analysis using three different enzyme digests to distinguish the eight alleles and 20 of the 32 haplotypes of the locus.

Length and sequence homology of primers

Each locus-specific primer includes a number of nucleotides which, under the conditions used in the hybridization, are sufficient to hybridize with an allele of the locus to be amplified and to be free from hybridization with alleles of other loci. The specificity of the primer increases with the number of nucleotides in its sequence under conditions that provide the same stringency. Therefore, longer primers are desirable. Sequences with fewer than 15 nucleotides are less certain to be specific for a particular locus. That is, sequences with fewer than 15 nucleotides are more likely to be present in a portion of the DNA associated with other genetic loci, particularly loci of other common origin or evolutionarily closely related origin, in inverse proportion to the length of the nucleotide sequence.

Each primer preferably includes at least about 15 nucleotides, more preferably at least about 20 nucleotides. The primer preferably does not exceed about 30 nucleotides, more preferably about 25 nucleotides. Most preferably, the primers have between about 20 and about 25 nucleotides.

A number of preferred primers are described herein. Each of those primers hybridizes with at least about 15 consecutive nucleotides of the designated region of the allele sequence. For many of the primers, the sequence is not identical for all of the other alleles of the locus. For each of the primers, additional preferred primers have sequences which correspond to the sequences of the homologous region of other alleles of the locus or to their complements.

When two sets of primer pairs are used sequentially, with the second primer pair amplifying the product of the first primer pair, the primers can be the same size as those used for the first amplification. However, smaller primers can be used in the second amplification and provide the requisite specificity. These smaller primers can be selected to be allele-specific, if desired. The primers of the second primer pair can have 15 or fewer, preferably 8 to 12, more preferably 8 to 10 nucleotides. When two sets of primer pairs are used to produce two amplified sequences, the second amplified DNA sequence is used in the subsequent analysis of genetic variation and must meet the requirements discussed previously for the amplified DNA sequence.

The primers preferably have a nucleotide sequence that is identical to a portion of the DNA sequence to be amplified or its complement. However, a primer having two nucleotides that differ from the target DNA sequence or its complement also can be used. Any nucleotides that are not identical to the sequence or its complement are preferably not located at the 3' end of the primer. The 3' end of the primer preferably has at least two, preferably three or more, nucleotides that are complementary to the sequence to which the primer binds. Any nucleotides that are not identical to the sequence to be amplified or its complement will preferably not be adjacent in the primer sequence. More preferably, noncomplementary nucleotides in the primer sequence will be separated by at least three, more preferably at least five, nucleotides. The primers should have a melting temperature (T_m) from about 55° to 75° C. Preferably the T_m is from about 60° C. to about 65° C. to facilitate stringent amplification conditions.

The primers can be prepared using a number of methods, such as, for example, the phosphotriester and phosphodiester methods or automated embodiments thereof. The phosphodiester and phosphotriester methods are described in Cruthers, *Science* 230:281–285 (1985); Brown et al, *Meth. Enzymol.*, 68:109 (1979); and Nrang et al, *Meth. Enzymol.*, 68:90 (1979). In one automated method, diethylphosphoramidites which can be synthesized as described by Beaucage et al, *Tetrahedron letters,* 22:1859–1962 (1981) are used as starting materials. A method for synthesizing primer oligonucleotide sequences on a modified solid support is described in U.S. Pat. No. 4,458,066. Each of the above references is incorporated herein by reference in its entirety.

Exemplary primer sequences for analysis of Class I and Class II HLA loci; bovine leukocyte antigens, and cystic fibrosis are described herein.

Amplification

The locus-specific primers are used in an amplification process to produce a sufficient amount of DNA for the analysis method. For production of RFLP fragment patterns or PDLP patterns which are analyzed by electrophoresis, about 1 to about 500 ng of DNA is required. A preferred amplification method is the polymerase chain reaction (PCR). PCR amplification methods are described in U.S. Pat. No. 4,683,195 (to Mullis et al, issued Jul. 28, 1987); U.S Pat. No. 4,683,194 (to Saiki et al, issued Jul. 28, 1987); Saiki et al, *Science,* 230:1350–1354 (1985); Scharf et al, *Science,* 324:163–166 (1986); Kogan et al, *New Engl. J. Med,* 317:985–990 (1987) and Saiki, Gyllensten and Erlich, *The Polymerase Chain Reaction in Genome Analysis: A Practical Approach,* ed. Davies pp. 141–152, (1988) I.R.L. Press, Oxford. Each of the above references is incorporated herein by reference in its entirety.

Prior to amplification, a sample of the individual organism's DNA is obtained. All nucleated cells contain genomic DNA and, therefore, are potential sources of the required

DNA. For higher animals, peripheral blood cells are typically used rather than tissue samples. As little as 0.01 to 0.05 cc of peripheral blood provides sufficient DNA for amplification. Hair, semen and tissue can also be used as samples. In the case of fetal analyses, placental cells or fetal cells present in amniotic fluid can be used. The DNA is isolated from nucleated cells under conditions that minimize DNA degradation. Typically, the isolation involves digesting the cells with a protease that does not attack DNA at a temperature and pH that reduces the likelihood of DNase activity. For peripheral blood cells, lysing the cells with a hypotonic solution (water) is sufficient to release the DNA.

DNA isolation from nucleated cells is described by Kan et al, *N. Engl. J. Med.* 297:1080–1084 (1977); Kan et al, *Nature* 251:392–392 (1974); and Kan et al, *PNAS* 75:5631–5635 (1978). Each of the above references is incorporated herein by reference in its entirety. Extraction procedures for samples such as blood, semen, hair follicles, semen, mucous membrane epithelium and other sources of genomic DNA are well known. For plant cells, digestion of the cells with cellulase releases DNA. Thereafter DNA is purified as described above.

The extracted DNA can be purified by dialysis, chromatography, or other known methods for purifying polynucleotides prior to amplification. Typically, the DNA is not purified prior to amplification.

The amplified DNA sequence is produced by using the portion of the DNA and its complement bounded by the primer pair as a template. As a first step in the method, the DNA strands are separated into single stranded DNA. This strand separation can be accomplished by a number of methods including physical or chemical means. A preferred method is the physical method of separating the strands by heating the DNA until it is substantially (approximately 93%) denatured. Heat denaturation involves temperatures ranging from about 80° to 105° C. for times ranging from about 15 to 30 seconds. Typically, heating the DNA to a temperature of from 90° to 93° C. for about 30 seconds to about 1 minute is sufficient.

The primer extension product(s) produced are complementary to the primer-defined region of the DNA and hybridize therewith to form a duplex of equal length strands. The duplexes of the extension products and their templates are then separated into single-stranded DNA. When the complementary strands of the duplexes are separated, the strands are ready to be used as a template for the next cycle of synthesis of additional DNA strands.

Each of the synthesis steps can be performed using conditions suitable for DNA amplification. Generally, the amplification step is performed in a buffered aqueous solution, preferably at a pH of about 7 to about 9, more preferably about pH 8. A suitable amplification buffer contains Tris-HCl as a buffering agent in the range of about 10 to 100 mM. The buffer also includes a monovalent salt, preferably at a concentration of at least about 10 mM and not greater than about 60 mM. Preferred monovalent salts are KCl, NaCl and $(NH_4)_2SO_4$. The buffer also contains $MgCl_2$ at about 5 to 50 mM. Other buffering systems such as hepes or glycine-NaOH and potassium phosphate buffers can be used. Typically, the total volume of the amplification reaction mixture is about 50 to 100 μl.

Preferably, for genomic DNA, a molar excess of about 10^6:1 primer:template of the primer pair is added to the buffer containing the separated DNA template strands. A large molar excess of the primers improves the efficiency of the amplification process. In general, about 100 to 150 ng of each primer is added.

The deoxyribonucleotide triphosphates dATP, dCTP, dGTP and dTTP are also added to the amplification mixture in amounts sufficient to produce the amplified DNA sequences. Preferably, the dNTPs are present at a concentration of about 0.75 to about 4.0 mM, more preferably about 2.0 mM. The resulting solution is heated to about 90° to 93° C. for from about 30 seconds to about 1 minute to separate the strands of the DNA. After this heating period the solution is cooled to the amplification temperature.

Following separation of the DNA strands, the primers are allowed to anneal to the strands. The annealing temperature varies with the length and GC content of the primers. Those variables are reflected in the T_m of each primer. Exemplary HLA DQA1 primers of this invention, described below, require temperatures of about 55° C. The exemplary HLA Class I primers of this invention require slightly higher temperatures of about 62° to about 68° C. The extension reaction step is performed following annealing of the primers to the genomic DNA.

An appropriate agent for inducing or catalyzing the primer extension reaction is added to the amplification mixture either before or after the strand separation (denaturation) step, depending on the stability of the agent under the denaturation conditions. The DNA synthesis reaction is allowed to occur under conditions which are well known in the art. This synthesis reaction (primer extension) can occur at from room temperature up to a temperature above which the polymerase no longer functions efficiently. Elevating the amplification temperature enhances the stringency of the reaction. As stated previously, stringent conditions are necessary to ensure that the amplified sequence and the DNA template sequence contain the same nucleotide sequence, since substitution of nucleotides can alter the restriction sites of probe binding sites in the amplified sequence.

The inducing agent may be any compound or system which facilitates synthesis of primer extension products, preferably enzymes. Suitable enzymes for this purpose include DNA polymerases (such as, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase), reverse transcriptase, and other enzymes (including heat-stable polymerases) which facilitate combination of the nucleotides in the proper manner to form the primer extension products. Most preferred is Taq polymerase or other heat-stable polymerases which facilitate DNA synthesis at elevated temperatures (about 60° to 90° C.). Taq polymerase is described, e.g., by Chien et al, *J. Bacteriol.*, 127:1550–1557 (1976). That article is incorporated herein by reference in its entirety. When the extension step is performed at about 72° C., about 1 minute is required for every 1000 bases of target DNA to be amplified.

The synthesis of the amplified sequence is initiated at the 3' end of each primer and proceeds toward the 5' end of the template along the template DNA strand, until synthesis terminates, producing DNA sequences of different lengths. The newly synthesized strand and its complementary strand form a double-stranded molecule which is used in the succeeding steps of the process. In the next step, the strands of the double-stranded molecule are separated (denatured) as described above to provide single-stranded molecules.

New DNA is synthesized on the single-stranded template molecules. Additional polymerase, nucleotides and primers can be added if necessary for the reaction to proceed under the conditions described above. After this step, half of the extension product consists of the amplified sequence bounded by the two primers. The steps of strand separation

and extension product synthesis can be repeated as many times as needed to produce the desired quantity of the amplified DNA sequence. The amount of the amplified sequence produced accumulates exponentially. Typically, about 25 to 30 cycles are sufficient to produce a suitable amount of the amplified DNA sequence for analysis.

The amplification method can be performed in a step-wise fashion where after each step new reagents are added, or simultaneously, where all reagents are added at the initial step, or partially step-wise and partially simultaneously, where fresh reagent is added after a given number of steps. The amplification reaction mixture can contain, in addition to the sample genomic DNA, the four nucleotides, the primer pair in molar excess, and the inducing agent, e.g., Taq polymerase.

Each step of the process occurs sequentially notwithstanding the initial presence of all the reagents. Additional materials may be added as necessary. Typically, the polymerase is not replenished when using a heat-stable polymerase. After the appropriate number of cycles to produce the desired amount of the amplified sequence, the reaction may be halted by inactivating the enzymes, separating the components of the reaction or stopping the thermal cycling.

In a preferred embodiment of the method, the amplification includes the use of a second primer pair to perform a second amplification following the first amplification. The second primer pair defines a DNA sequence which is a portion of the first amplified sequence. That is, at least one of the primers of the second primer pair defines one end of the second amplified sequence which is within the ends of the first amplified sequence. In this way, the use of the second primer pair helps to ensure that any amplified sequence produced in the second amplification reaction is specific for the tested locus. That is, non-target sequences which may be copied by a locus-specific pair are unlikely to contain sequences that hybridize with a second locus-specific primer pair located within the first amplified sequence.

In another embodiment, the second primer pair is specific for one allele of the locus. In this way, detection of the presence of a second amplified sequence indicates that the allele is present in the sample. The presence of a second amplified sequence can be determined by quantitating the amount of DNA at the start and the end of the second amplification reaction. Methods for quantitating DNA are well known and include determining the optical density at 260 (OD_{260}), and preferably additionally determining the ratio of the optical density at 260 to the optical density at 280 (OD_{260}/OD_{280}) to determine the amount of DNA in comparison to protein in the sample.

Preferably, the first amplification will contain sufficient primer for only a limited number of primer extension cycles, e.g. less than 15, preferably about 10 to 12 cycles, so that the amount of amplified sequence produced by the process is sufficient for the second amplification but does not interfere with a determination of whether amplification occurred with the second primer pair. Alternatively, the amplification reaction can be continued for additional cycles and aliquoted to provide appropriate amounts of DNA for one or more second amplification reactions. Approximately 100 to 150 ng of each primer of the second primer pair is added to the amplification reaction mixture. The second set of primers is preferably added following the initial cycles with the first primer pair. The amount of the first primer pair can be limited in comparison to the second primer pair so that, following addition of the second pair, substantially all of the amplified sequences will be produced by the second pair.

As stated previously, the DNA can be quantitated to determine whether an amplified sequence was produced in the second amplification. If protein in the reaction mixture interferes with the quantitation (usually due to the presence of the polymerase), the reaction mixture can be purified, as by using a 100,000 MW cut off filter. Such filters are commercially available from Millipore and from Centricon.

Analysis of the Amplified DNA Sequence

As discussed previously, the method used to analyze the amplified DNA sequence to characterize the allele(s) present in the sample DNA depends on the genetic variation in the sequence. When distinctions between alleles include primer-defined length polymorphisms, the amplified sequences are separated based on length, preferably using gel or capillary electrophoresis. When using probe hybridization for analysis, the amplified sequences are reacted with labeled probes. When the analysis is based on RFLP fragment patterns, the amplified sequences are digested with one or more restriction endonucleases to produce a digest and the resultant fragments are separated based on length, preferably using gel or capillary electrophoresis. When the only variation encompassed by the amplified sequence is a sequence variation that does not result in a change in length or a change in a restriction site and is unsuitable for detection by a probe, the amplified DNA sequences are sequenced.

Procedures for each step of the various analytical methods are well known and are described below.

Production of RFLP Fragment Patterns

Restriction endonucleases

A restriction endonuclease is an enzyme that cleaves or cuts DNA hydrolytically at a specific nucleotide sequence called a restriction site. Endonucleases that produce blunt end DNA fragments (hydrolysis of the phosphodiester bonds on both DNA strands occur at the same site) as well as endonucleases that produce sticky ended fragments (the hydrolysis sites on the strands are separated by a few nucleotides from each other) can be used.

Restriction enzymes are available commercially from a number of sources including Sigma Pharmaceuticals, Bethesda Research Labs, Boehringer-Manheim and Pharmacia. As stated previously, a restriction endonuclease used in the present invention cleaves an amplified DNA sequence of this invention to produce a digest comprising a set of fragments having distinctive fragment lengths. In particular, the fragments for one allele of a locus differ in size from the fragments for other alleles of the locus. The patterns produced by separation and visualization of the fragments of a plurality of digests are sufficient to distinguish each allele of the locus. More particularly, the endonucleases are chosen so that by using a plurality of digests of the amplified sequence, preferably fewer than five, more preferably two or three digests, the alleles of a locus can be distinguished.

In selecting an endonuclease, the important consideration is the number of fragments produced for amplified sequences of the various alleles of a locus. More particularly, a sufficient number of fragments must be produced to distinguish between the alleles and, if required, to provide for individuality determinations. However, the number of fragments must not be so large or so similar in size that a pattern that is not distinguishable from those of other haplotypes by the particular detection method is produced. Preferably, the fragments are of distinctive sizes for each

allele. That is, for each endonuclease digest of a particular amplified sequence, the fragments for an allele preferably differ from the fragments for every other allele of the locus by at least 10, preferably 20, more preferably 30, most preferably 50 or more nucleotides.

One of ordinary skill can readily determine whether an endonuclease produces RFLP fragments having distinctive fragment lengths. The determination can be made experimentally by cleaving an amplified sequence for each allele with the designated endonuclease in the invention method. The fragment patterns can then be analyzed. Distinguishable patterns will be readily recognized by determining whether comparison of two or more digest patterns is sufficient to demonstrate characteristic differences between the patterns of the alleles.

The number of digests that need to be prepared for any particular analysis will depend on the desired information and the particular sample to be analyzed. Since HLA analyses are used for a variety of purposes ranging from individuality determinations for forensics and paternity to tissue typing for transplantation, the HLA complex will be used as exemplary.

A single digest may be sufficient to determine that an individual cannot be the person whose blood was found at a crime scene. In general, however, where the DNA samples do not differ, the use of two to three digests for each of two to three HLA loci will be sufficient for matching applications (forensics, paternity). For complete HLA typing, each locus needs to be determined.

In a preferred embodiment, sample HLA DNA sequences are divided into aliquots containing similar amounts of DNA per aliquot and are amplified with primer pairs (or combinations of primer pairs) to produce amplified DNA sequences for a number of HLA loci. Each amplification mixture contains only primer pairs for one HLA locus. The amplified sequences are preferably processed concurrently, so that a number of digest RFLP fragment patterns can be produced from one sample. In this way, the HLA type for a number of alleles can be determined simultaneously.

Alternatively, preparation of a number of RFLP fragment patterns provides additional comparisons of patterns to distinguish samples for forensic and paternity analyses where analysis of one locus frequently fails to provide sufficient information for the determination when the sample DNA has the same allele as the DNA to which it is compared.

Production of RFLP fragments

Following amplification, the amplified DNA sequence is combined with an endonuclease that cleaves or cuts the amplified DNA sequence hydrolytically at a specific restriction site. The combination of the endonuclease with the amplified DNA sequence produces a digest containing a set of fragments having distinctive fragment lengths. U.S. Pat. No. 4,582,788 (to Erlich, issued Apr. 15, 1986) describes an HLA typing method based on restriction length polymorphism (RFLP). That patent is incorporated herein by reference in its entirety.

In a preferred embodiment, two or more aliquots of the amplification reaction mixture having approximately equal amounts of DNA per aliquot are prepared. Conveniently about 5 to about 10 μl of a 100 μl reaction mixture is used for each aliquot. Each aliquot is combined with a different endonuclease to produce a plurality of digests. In this way, by using a number of endonucleases for a particular amplified DNA sequence, locus-specific combinations of endonucleases that distinguish a plurality of alleles of a particular locus can be readily determined. Following preparation of the digests, each of the digests can be used to form RFLP patterns. Preferably, two or more digests can be pooled prior to pattern formation.

Alternatively, two or more restriction endonucleases can be used to produce a single digest. The digest differs from one where each enzyme is used separately and the resultant fragments are pooled since fragments produced by one enzyme may include one or more restriction sites recognized by another enzyme in the digest. Patterns produced by simultaneous digestion by two or more enzymes will include more fragments than pooled products of separate digestions using those enzymes and will be more complex to analyze.

Furthermore, one or more restriction endonucleases can be used to digest two or more amplified DNA sequences. That is, for more complete resolution of all the alleles of a locus, it may be desirable to produce amplified DNA sequences encompassing two different regions. The amplified DNA sequences can be combined and digested with at least one restriction endonuclease to produce RFLP patterns.

The digestion of the amplified DNA sequence with the endonuclease can be carried out in an aqueous solution under conditions favoring endonuclease activity. Typically the solution is buffered to a pH of about 6.5 to 8.0. Mild temperatures, preferably about 20° C. to about 45° C., more preferably physiological temperatures (25° to 40° C.), are employed. Restriction endonucleases normally require magnesium ions and, in some instances, cofactors (ATP and S-adenosyl methionine) or other agents for their activity. Therefore, a source of such ions, for instance inorganic magnesium salts, and other agents, when required, are present in the digestion mixture. Suitable conditions are described by the manufacturer of the endonuclease and generally vary as to whether the endonuclease requires high, medium or low salt conditions for optimal activity.

The amount of DNA in the digestion mixture is typically in the range of 1% to 20% by weight. In most instances 5 to 20 μg of total DNA digested to completion provides an adequate sample for production of RFLP fragments. Excess endonuclease, preferably one to five units/μg DNA, is used.

The set of fragments in the digest is preferably further processed to produce RFLP patterns which are analyzed. If desired, the digest can be purified by precipitation and resuspension as described by Kan et al, *PNAS* 75:5631–5635 (1978), prior to additional processing. That article is incorporated herein by reference in its entirety.

Once produced, the fragments are analyzed by well known methods. Preferably, the fragments are analyzed using electrophoresis. Gel electrophoresis methods are described in detail hereinafter. Capillary electrophoresis methods can be automated (as by using Model 207A analytical capillary electrophoresis system from Applied Biosystems of Foster City, Calif.) and are described in Chin et al, *American Biotechnology Laboratory News Edition,* December, 1989.

Electrophoretic Separation of DNA Fragments

Electrophoresis is the separation of DNA sequence fragments contained in a supporting medium by size and charge under the influence of an applied electric field. Gel sheets or slabs, e.g. agarose, agarose-acrylamide or polyacrylamide, are typically used for nucleotide sizing gels. The electrophoresis conditions affect the desired degree of resolution of

the fragments. A degree of resolution that separates fragments that differ in size from one another by as little as 10 nucleotides is usually sufficient. Preferably, the gels will be capable of resolving fragments which differ by 3 to 5 nucleotides. However, for some purposes (where the differences in sequence length are large), discrimination of sequence differences of at least 100 nt may be sufficiently sensitive for the analysis.

Preparation and staining of analytical gels is well known. For example, a 3% Nusieve 1% agarose gel which is stained using ethidium bromide is described in Boerwinkle et al, *PNAS,* 86:212–216 (1989). Detection of DNA in polyacrylamide gels using silver stain is described in Goldman et al, *Electrophoresis,* 3:24–26 (1982); Marshall, *Electrophoresis,* 4:269–272 (1983); Tegelstrom, *Electrophoresis,* 7:226–229 (1987); and Allen et al, *BioTechniques* 7:736–744 (1989). The method described by Allen et al, using large-pore size ultrathin-layer, rehydratable polyacrylamide gels stained with silver is preferred. Each of those articles is incorporated herein by reference in its entirety.

Size markers can be run on the same gel to permit estimation of the size of the restriction fragments. Comparison to one or more control sample(s) can be made in addition to or in place of the use of size markers. The size markers or control samples are usually run in one or both the lanes at the edge of the gel, and preferably, also in at least one central lane. In carrying out the electrophoresis, the DNA fragments are loaded onto one end of the gel slab (commonly called the "origin") and the fragments separate by electrically facilitated transport through the gel, with the shortest fragment electrophoresing from the origin towards the other (anode) end of the slab at the fastest rate. An agarose slab gel is typically electrophoresed using about 100 volts for 30 to 45 minutes. A polyacrylamide slab gel is typically electrophoresed using about 200 to 1,200 volts for 45 to 60 minutes.

After electrophoresis, the gel is readied for visualization. The DNA fragments can be visualized by staining the gel with a nucleic acid-specific stain such as ethidium bromide or, preferably, with silver stain, which is not specific for DNA. Ethidium bromide staining is described in Boerwinkle et al, supra. Silver staining is described in Goldman et al, supra, Marshall, supra, Tegelstrom, supra, and Allen et al, supra.

Probes

Allele-specific oligonucleotides or probes are used to identify DNA sequences which have regions that hybridize with the probe sequence. The amplified DNA sequences defined by a locus-specific primer pair can be used as probes in RFLP analyses using genomic DNA. U.S. Pat. No. 4,582,788 (to Erlich, issued Apr. 15, 1986) describes an exemplary HLA typing method based on analysis of RFLP patterns produced by genomic DNA. The analysis uses cDNA probes to analyze separated DNA fragments in a Southern blot type of analysis. As stated in the patent "[C]omplementary DNA probes that are specific to one (locus-specific) or more (multilocus) particular HLA DNA sequences involved in the polymorphism are essential components of the hybridization step of the typing method" (col. 6, l. 3–7).

The amplified DNA sequences of the present method can be used as probes in the method described in that patent or in the present method to detect the presence of an amplified DNA sequence of a particular allele. More specifically, an amplified DNA sequence having a known allele can be produced and used as a probe to detect the presence of the allele in sample DNA which is amplified by the present method.

Preferably, however, when a probe is used to distinguish alleles in the amplified DNA sequences of the present invention, the probe has a relatively short sequence (in comparison to the length of the amplified DNA sequence) which minimizes the sequence homology of other alleles of the locus with the probe sequence. That is, the probes will correspond to a region of the amplified DNA sequence which has the largest number of nucleotide differences from the amplified DNA sequences of other alleles produced using that primer pair.

The probes can be labelled with a detectable atom, radical or ligand using known labeling techniques. Radiolabels, usually ^{32}P, are typically used. The probes can be labeled with ^{32}P by nick translation with an α-^{32}P-dNTP (Rigby et al, *J. Mol. Biol.,* 113:237 (1977)) or other available procedures to make the locus-specific probes for use in the methods described in the patent. The probes are preferably labeled with an enzyme, such as hydrogen peroxidase. Coupling enzyme labels to nucleotide sequences are well known. Each of the above references is incorporated herein by reference in its entirety.

The analysis method known as "Southern blotting" that is described by Southern, *J. Mol. Biol.,* 98:503–517 (1975) is an analysis method that relies on the use of probes. In Southern blotting the DNA fragments are electrophoresed, transferred and affixed to a support that binds nucleic acid, and hybridized with an appropriately labeled cDNA probe. Labeled hybrids are detected by autoradiography, or preferably, use of enzyme labels.

Reagents and conditions for blotting are described by Southern, supra; Wahl et al, *PNAS* 6:3683–3687 (1979); Kan et al, *PNAS, supra, U.S. Pat. No.* 4,302,204 and *Molecular Cloning: A Laboratory Manual* by Maniatis et al, Cold Spring Harbor Laboratory 1982. After the transfer is complete the paper is separated from the gel and is dried. Hybridization (annealing) of the resolved single stranded DNA on the paper to an probe is effected by incubating the paper with the probe under hybridizing conditions. See Southern, supra; Kan et al, *PNAS,* supra and U.S. Pat. No. 4,302,204, col 5, line 8 et seq. Complementary DNA probes specific for one allele, one locus (locus-specific) or more are essential components of the hybridization step of the typing method. Locus-specific probes can be made by the amplification method for locus-specific amplified sequences, described above. The probes are made detectable by labeling as described above.

The final step in the Southern blotting method is identifying labeled hybrids on the paper (or gel in the solution hybridization embodiment). Autoradiography can be used to detect radiolabel-containing hybrids. Enzyme labels are detected by use of a color development system specific for the enzyme. In general, the enzyme cleaves a substrate, which cleavage either causes the substrate to develop or change color. The color can be visually perceptible in natural light or a fluorochrome which is excited by a known wavelength of light.

Sequencing

Genetic variations in amplified DNA sequences which reflect allelic difference in the sample DNA can also be detected by sequencing the amplified DNA sequences.

Methods for sequencing oligonucleotide sequences are well known and are described in, for example, *Molecular Cloning: A Laboratory Manual* by Maniatis et al, Cold Spring Harbor Laboratory 1982. Currently, sequencing can be automated using a number of commercially available instruments.

Due to the amount of time currently required to obtain sequencing information, other analysis methods, such as gel electrophoresis of the amplified DNA sequences or a restriction endonuclease digest thereof are preferred for clinical analyses.

Kits

As stated previously, the kits of this invention comprise one or more of the reagents used in the above described methods. In one embodiment, a kit comprises at least one genetic locus-specific primer pair in a suitable container. Preferably the kit contains two or more locus-specific primer pairs. In one embodiment, the primer pairs are for different loci and are in separate containers. In another embodiment, the primer pairs are specific for the same locus. In that embodiment, the primer pairs will preferably be in the same container when specific for different alleles of the same genetic locus and in different containers when specific for different portions of the same allele sequence. Sets of primer pairs which are used sequentially can be provided in separate containers in one kit. The primers of each pair can be in separate containers, particularly when one primer is used in each set of primer pairs. However, each pair is preferably provided at a concentration which facilitates use of the primers at the concentrations required for all amplifications in which it will be used.

The primers can be provided in a small volume (e.g. 100 μl) of a suitable solution such as sterile water or Tris buffer and can be frozen. Alternatively, the primers can be air dried.

In another embodiment, a kit comprises, in separate containers, two or more endonucleases useful in the methods of this invention. The kit will preferably contain a locus-specific combination of endonucleases. The endonucleases can be provided in a suitable solution such as normal saline or physiologic buffer with 50% glycerol (at about −20° C.) to maintain enzymatic activity.

The kit can contain one or more locus-specific primer pairs together with locus-specific combinations of endonucleases and may additionally include a control. The control can be an amplified DNA sequence defined by a locus-specific primer pair or DNA having a known HLA type for a locus of interest.

Additional reagents such as amplification buffer, digestion buffer, a DNA polymerase and nucleotide triphosphates can be provided separately or in the kit. The kit may additionally contain gel preparation and staining reagents or preformed gels.

Analyses of exemplary genetic loci are described below.

Analysis of HLA Type

The present method of analysis of genetic variation in an amplified DNA sequence to determine allelic difference in sample DNA can be used to determine HLA type. Primer pairs that specifically amplify genomic DNA associated with one HLA locus are described in detail hereinafter. In a preferred embodiment, the primers define a DNA sequence that contains all exons that encode allelic variability associated with the HLA locus together with at least a portion of one of the adjacent intron sequences. For Class I loci, the variable exons are the second and third exons. For Class II loci, the variable exon is the second exon. The primers are preferably located so that a substantial portion of the amplified sequence corresponds to intron sequences.

The intron sequences provide restriction sites that, in comparison to cDNA sequences, provide additional information about the individual; e.g., the haplotype. Inclusion of exons within the amplified DNA sequences does not provide as many genetic variations that enable distinction between alleles as an intron sequence of the same length, particularly for constant exons. This additional intron sequence information is particularly valuable in paternity determinations and in forensic applications. It is also valuable in typing for transplant matching in that the variable lengths of intron sequences included in the amplified sequence produced by the primers enables a distinction to be made between certain heterozygotes (two different alleles) and homozygotes (two copies of one allele).

Allelic differences in the DNA sequences of HLA loci are illustrated below. The tables illustrate the sequence homology of various alleles and indicate exemplary primer binding sites. Table 1 is an illustration of the alignment of the nucleotides of the Class I A2, A3, Ax, A24 (formerly referred to as A9), B27, B58 (formerly referred to as B17), C1, C2 and C3 allele sequences in intervening sequence (IVS) I and III. (The gene sequences and their numbering that are used in the tables and throughout the specification can be found in the Genbank and/or European Molecular Biology Laboratories (EMBL) sequence databanks. Those sequences are incorporated herein by reference in their entirety.) Underlined nucleotides represent the regions of the sequence to which exemplary locus-specific or Class I-specific primers bind.

Table 2 illustrates the alignment of the nucleotides in IVS I and II of the DQA3 (now DQA1 0301), DQA1.2 (now DQA1 0102) and DQA4.1 (now DQA1 0501) alleles of the DQA1 locus (formerly referred to as the DR4, DR6 and DR3 alleles of the DQA1 locus, respectively). Underlined nucleotides represent the regions of the sequence to which exemplary DQA1 locus-specific primers bind.

Table 3 illustrates the alignment of the nucleotides in IVS I, exon 2 and IVS II of two individuals having the $DQw1_v$ allele (designated hereinafter as $DQw1_v$a and $DQw1_v$b for the upper and lower sequences in the table, respectively), the DQw2 and DQw8 alleles of the DQB1 locus. Nucleotides indicated in the $DQw1_v$b, DQw2 and DQw8 allele sequences are those which differ from the $DQw1_v$a sequence. Exon 2 begins and ends at nt 599 and nt 870 of the $DQw1_v$a allele sequence, respectively. Underlined nucleotides represent the regions of the sequence to which exemplary DQB1 locus-specific primers bind.

Table 4 illustrates the alignment of the nucleotides in IVS I, exon 2 and IVS II of the DPB4.1, DPB9, New and DPw3 alleles of the DPB1 locus. Nucleotides indicated in the DPB9, New and DPw3 allele sequences are those which differ from the DPB4.1 sequence. Exon 2 begins and ends at nt 7644 and nt 7907 of the DPB4.1 allele sequence, respectively. Underlined nucleotides represent the regions of the sequence to which exemplary DPB1 locus-specific primers bind.

TABLE 1

```
Class I Seq

C1     1                    GATTACCAATATTGTGCGACCTACTGTATCAATAAAC
C2     1                                      T

C1    38 AAAAAGGAAACTGGTCTCTATGAGAATCTCTACCTGGTGCTTTCAGACAA
C2    38                    G G

C1    88 CACTTCACCAGGTTTAAAGAGAAAACTCCTGACTCTACACGTCCATTCCC
C2    88

B27    1       GAGCTCACTCTCTGGCATCAAGTTC                TCCGTG
C1   138 AGGGCGAGCTCACTGTCTGGCAGCAAGTTCCCCATGGTCGAGTTTCCCTG
C2   138                          T                 -

A2     1     AAGCTTACTCTCTGGCACCAAAC  TCCATGGGATGATTTTTCCTTCC TAG
B27   32                                        ATCAGTTTCCCT
C1   188 TACAAGAGTCCAAGGGGAGAGGTAAGTGTCCTTT  AT    TTTGCTGGATGTAG
C2   187

A2    50      AAGAGTCCAGGTGGACAGGTAA GGAGTGGGAGT        CAGGGAGTC
B27   44 ACACAAGA TCCAAGAGGAGAGGTAA GGAGT  GAG      AGGCAGGGAGTC
C1   238 TTTAATATTACCT GAGGTAAGGTAA GGC AAAGAGTGGG AGGCAGGGAGTC
C2   237                                C  -            G

A2    98 CAGTTCCAGGGACAGAGATTACGGGATAAAAAGTGAAAGGAGAGGGACG  GGGCCCAT
B27   91 CAGTT CAGGGACAGGGATTCCAGGAGGAGAAGTGAAGGGGAAGC GGG TGGGC
C1   288 CAGTT CAGGGACGGGGATTCCAGGAGAAG   TGAAGGGGAAG  GGGCTGGGCG
C2   288                                                    -

A2   149   GCCGAG   GGTTTCTCCCTTGTTTCT CAGACAGCTC TTGGGCCA A GAC
B27  141   GCCACTGGGGGTCTCTCCCTGGTTTCCACAGACAGATCCTTGTGCC   GGAC
C1   338 CAGCC  TGGGGGTCTCTCCCTGGTTTCCACAGACAGATCCTTG GCC  AGGAC
C2   337                                              - -  GG

A2   195 TCAGGGAGACATTGAGACAGAGC GCTTGGCACAGAAGCAGAGGGGTCAGGG
B27  191 TCAGGCAGACAGTGTGACAAAGAGGCT GGTGTAGGAGAAGAGGGATCAGG
C1   388 TCAGGCACACAGTGTGACAAAGATGCTTGGTGTAGGAGAAGAGGGATCAG
C2   387                                                 G

A2   246  CGAA GTCCAGGGCCCCAGGCGTTGGCTCTCAGGGTCTCAGGCCCCGAAGG
A3     1
Ax     1
A24    1                                                  -
B27  241 ACGAACGTCCAAGGCCCCGGGCG CGG TCTCAGGGTCTCAGGCTCCGAGAG
C1   438 ACGAA GTCCCAGGTCCCGGGCG GGGTTCTCAGGGTCTCAGGCTCCAAGGG
C2   438               -A

A2   296 CGGTGTATGGATTGGGGAGTCCCAGCCTTGGGGATTCCCCAACTCCGC AGTT
A3     9  T      A                    -
Ax     9                     TG                 G   C
A24   11                                        -     -  T
B27  291 CCTTGTCTGCATTGGGGAGGCGCACAGTTGGGG TTCCCCACTCCCACGAGTT
C1   488 GCGTGTCTGCACTGGGGAGGCGCCGCGTTGAGGATTCTCCACTCCCCTGA
C2   488

A2   348 TCTTTTCTCCC  TCTCCCAACCTATGTAGGGTCCTTCTTCCTGGAT ACTCAC
A3    60             CTG             C               A                 G
Ax    61    C     ---        A       GC AC                C
A24   61               TG-                             -
B27  344 TCACTTCT      TCTCCCAACCTATGTCGGGTCCTTCTTCCAGGAT ACTCGT
C1   538   G TTCACTTCTTCTCCCAACCTGCGTCGGGTCCTTCTTCCTGAAT ACTCAT
C2   538    T                          A
C3     1                        T  G                        G

A2   399 GACGCGGACCCAGTTCTCACTCCCATTGGGTGTCGGGTTTCC    AGAGAAG C
A3   114
Ax   109   A       A          T      C A                -  T
A24  111                                              G
B27  392 GACGCGTCCCCATTTC CACTCCCATTGGGTGTCGGGT    GTCTAGAGAAG C
B58    1
C1   588 GACGCGTCCCCAATTCCCACTCCCATTGGGTGTCGGGT     TCT  AGAAG C
C2   589                          -                      AG
C3    36                                      -ACCNN            G
```

TABLE 1-continued

```
A2   449 CAATCAGTGTCGTCGCGGTCGCGGTTCTAAAGT CCGCACG
A3   164                          T          C
Ax   159       G    C  C         C                C
A24  161        A                  T
B27  442 CAATCAGTGTCGCCGGGGTCCCAGTTCTAAAGT CCCCACG
B58   12
C1   635 CAATCAGCGTCTCCGCAGTCCCGGTTCTAAAGTCCC CAGT
C2   637        C
C3    87  GG                          G

A2   489 CACCCACCGGGACTCAGA TTCTCCCCAGACGCCGAGGATGGC                C
A3   204                                                 TCGTGGAGACCAGGC
Ax   199                                                T               G
A24  201
B27  482 CACCCACCCGGACTCAGA ATCTCCTCAGACGCCGAG ATGCG                  G
B58   52
C1   675 CACCCACCCGGACTCAGA TTCTCCCCAGACGCCGAG ATGCG                  G
C2   677                   G
C3   127

          1st EXON
A2   530 GTCATGGCGCCCCGAACCCTCGTCCTGCTACTCTCGGGGGCTC
A3   262                          C                    C
Ax   242 C                        C         G      A   C
A24  244          G                                    C
B27  524 GTCACGGCGCCCCGAACCCTCCTCCTGCTGCTCTGGGGGGCAG
B58   94                      G
C1   717 GTCATGGCGCCCCGAACCCTCATCCTGCTGCTCTCGGGAGCCC
C2   719
C3   169                   G

A2   574 TGGCCCTGACCCAGACCTGGGCGG
A3   305
Ax   285                            C
A24  287                           A
B27  567 TGGCCCTGACCGAGACCTGGGCTG
B58  137                           C
C2   760 TGGCCCTGACCGAGACCTGGGCCT
C2   762
C3   212                            G

          IVS1
A2   599 GTGAGTGCGGGGTCGGG AGGGAAACG GCC TCTGT GGGGAGAAGCAACGGGCC G
A3   329                            C AC          C                G       T
Ax   309          A      T C          T-G --    --- -      G  NG G      CG
A24  311                            TCG    C     C                G      CG
B27  591 GTGAGTGCGGGGTCAGGCAGGGAAATG GCC TCTGT GGGGAGGAGCGAGGGGA CG
B58  161                   G  -                                        C
C1   784 GTGAGTGCGGGGTTGGG AGGGAAACG GCC TCT GCGGAGAGGAACGAGGTGCCCG
C2   786                                                      G      G
C3   236                            T           T            G      G

A2   652 CCTGGC GGGGGCGCAGGACCCGGGAAGCCGCGCCGGGAGGAGGGTCGGGCGGGTCTCAG
A3   383                          G   G              C
Ax   357   C    G    T             A   G         A
A24  367                    A
B27  645  CAGGC GGGGGCGCAGGACCCGGGGAGCCGCGCCGGGAGGAGGGTCGGGCGGGTCTCAG
B58  215                          T A
C1   838 CCCGGC  AGG CGCAGGACCCGGGGAGCCGCGCAGGGAGGAGGGTCGGGCGGGTCTCAG
C2   840        G     G -             AGC
C3   291        GGA   G

A2   711 CCACTCCTCGTCCCCAG
A3   442       G    -C
Ax   417  TC         CT
A24  426
B27  703 CCCCTCCTCGCCCCCAG
B58  273
C1   895 CCCCTCCTCGCCCCCAG
C2   898           T
C3   351            -
```

TABLE 1-continued

```
          IVS3
A2   1515 GTACCAGGGGCCACGGGGCGCCTCCCTGATCGCCTGTAGATCTCCCGGGCTGGCCTCCC
A3   1245                  -
Ax   1222            C ACA  -
A24  1228                                                  G
B27  1508 GTACCAGGGGCAGTGGGGAGCCTTCCCCATCTCCTATAGGTCGCCGGGGATGGCCTCCC
B58  1082
C1   1704 GTACCAGGGGCAGTGGGGAGCCTTCCCCATCTCCCGTAGATCTCCCGGCATGGCCTCCC
C2   1705                                        T              G
C3   1155                  -                     T              G

A2   1574  ACAAGGAGGGGAGACAATTGGGACCAACACTAGAATATCGCCCTCCCTCTGGT
A3   1303                    C          C        G      A    T   T
Ax   1280        A A             A                T
A24  1287 C
B27  1567  ACGAGAAGAGGAGGAAAATGGGATCAGCGCTAGAATGTCGCCCTCCCTTGAAT
B58  1141
C1   1763  ACGAGGAGGGGAGGAAAATGGGATCAGCGCTAGAATATCGCCCTCCCTGAAAT
C2   1764
C3   1213

A2   1627 CCTGAGGGAGAGGAATCCTCCTGGGTTTCCAGATCCTGTACCAGAGAGTGA
A3   1356 T                  T   T   T       -  GA     G
Ax   1333  T                      T          ----------
A24  1341  T
B27  1620 GGAGAATGGCATGAGTTTTCCTGAGTTTC
B58  1194
C1   1816 GGAGAATGGGATGAGTTTTCCTGAGTTTC
C2   1817
C3   1266

A2   1678 CTCTGAGGTTCCGCCCTGCTCTCTGA CACAATTAAGGGATAAAATCTCTGAAGGA
A3   1406            T   G         A A -G                    -
Ax   1372          G     -                                 G      G  -
A24  1392                                                            C
B27  1649 CTCTGAGGGCCCCCTCTTCTCTCT AGGACAATTAAGGGATGACGTCTCTGAGGAA
B58  1223
C1   1845 CTCTGAGGGCCCCCTCTGCTCTCT AGGACAATTAAGGGATGAAGTCCTTGAGGAA
C2   1846
C3   1295                              G                              A

A2   1733 ATGACGGG AAGACGATCCCTCGAATACTGATGAGTGGTTCCCTTTGACAC
A3   1460 G                    T    T G  T    G                  G
Ax   1426  ATGAA   G       A       G
A24  1447          A                              C
B27  1704 ATGGAGGGGAAGACAGTCCCTAGAATACTGATCAGGGGTCCCCTTTGACCC
B58  1278
C1   1900 ATGGAGGGGAAGACAGTCCCTGGAATACTGATCAGGGGTCCCCTTTGACCA
C2   1901
C3   1351                          A

A2   1783       ACACAGGCAGCAGCCTTGGG CCCG    TGACTTTTCCTCTCAGGCCTTGTTCTCTGC
A3   1510       ------C    GA G
Ax   1477       ------T                 C
A24  1497       ------C                   A
B27  1755            CTGCAGCAGCCTTGGGAACCG    TGACTTTTCCTCTCAGGCCTTGTTCACAGC
B58  1329                                                          T T
C1   1951 CTTTGACCACTGCAGCAGCTGTGGTCAGGCTGCTGACCTTT CTCTCAGGCCTTGTTCTCTGC
C2   1952
C3   1411 ----------

A2   1837 TTCACACTCAATGTGTGTGGGGGTCTGAGTCCAGCACTTCTGAGTCCTTCAGCC
A3   1560                                                    C
Ax   1528                    C                  ----------C
A24  1547                                                    C
B27  1806 CTCACACTCAGTGTGTTTGGGGCTCTGATTCCAGCACTTCTGAGTCACTTTACC
B58  1380
C1   2013 CTCACGTTCAATGTGTTTGAAGGTTTGATTCCAGCTTTTCTGAGTCCTTCGGCC
C2   2014
C3   1464         C
```

TABLE 1-continued

```
A2   1891 TCCACTCAGGTCAGGACCAGAAGTCGCTGTTCCCTCTTCAGGGACTAGAA TTTCCACGGAATAG
A3   1614                                          TC        A
Ax   1567                                                    T
A24  1600                                                  A
B27  1860 TCCACTCAGATCAGGAGCAGAAGTCCCTGTTCCCCGCTCAGAGACT CGAACTTTCCAATGAATAG
B58  1434
C1   2067 TCCACTCAGGTCAGGACCAGAAGTCGCTGTTCCTCCCTCAGAGACTAGAACTTTCCAATGAATAG
C2   2068
C3   1518

A2   1955 GAGATTATCCCAGGTGCCTGTGTCCAGGCTGGTGTCTGGGTTCTGTGCTCCCTTCCCCA
A3   1664 —
Ax   1632          T T          C     T          T
A24  1650 —                 -              A                        A    T        G
B27  1925 GAGATTATCCCAGGTGCCTGCGTCCAGGCTGGTGTCTGGGTTCTGTGCCC CTTCCCCA
B58  1499
C1   2132 GAGATTATCCCAGGTGCCTGTGTCCAGGCTGGCGTCTGGGTTCTGTGCCCCCTTCCCCA
C2   2133
C3   1583

A2   2014 TCCCAGGTGTCCTGTCCATTCTCAAGA TAGCCACATGTGTGCTGGAGGAGTGTCCCATG
A3   1721      G                          G          C         T
Ax   1691 C   T    CA         A          G          C         T
A24  1706                                G          CA        T
B27  1983 CCCCAGGTGTCCTGTCCATTCTC AGGCTGGTCACATGGGTGGTCCTAGGGTGTCCCATG
B58  1557  A
C1   2191 CCCCAGGTGTCCTGTCCATTCTC AGGATGGTCACATGGGCGCTGTTGGAGTGTCGCAAG
C3   2192                                A
C3   1642                      G

A2   2073 ACAGATCGAAAATGCCTGAATGATCTGACTCT   TCCTGACAG      2113
A3   1780        GC                TT                C T         1820
Ax   1750        GC                TT          TT    C T         1791
A24  1765      G GCAAAA ————————————————————  -   C T         1784
B27  2042 AGAGATGCAAAGCGCCTGAATTTTCTGACTCTTCCCAT   CAG      2083
B58  1616                                                     1656
C1   2250 AGAGATACAAAGTGTCTGAATTTTCTGACTCTTCCCGT   CAG      2290
C2   2251                                          G        2292
C3   1701                                                   1741
```

TABLE 2

```
DQA1 Seq

A3   1 GATCTCTGTGTAGAATGTCCTGTTCTGAGCCAGTCCTGAGAGGAAAGGAAGTATAATCAA
A1.2 1              G       A
A4.1 1   C          G                                  A  A  C     G

A3   61 TTTGTTATTAACTGATGAAAGAATTAAGTGAAAGATAAACCTTAGGAAGC AGAGGGAAGT
A1.2 61             CA                              T  C        C
A4.1 61                                          G  T          C   A

A3   121 TAA      TCTATGACTAAGAAAGTTAAGTACTCTGATAACTCATTCATTCCTTCT
A1.2 122 A  CCTAA T C              C   A    A
A4.1 122 A  CCTAA   C              C   A   CA A

A3   172 TTTGTTCATTTACATT ATTTAATCACAAGTCTATGATGTGCCAGGCTCTCAGGAAATA
A1.2 178         A        \        T     C    C         A
A4.1 178         A       G         T     CG             A

A3   230 GTGAAAATTGG CACGCGATATTCTGCCCTTGTGTAGCACACACCGTAGTGGGAAAG
A1.2 236  A         A  T                          G      TAG
A4.1 237  A    C    A  T T                        G     TTA

A3   286 AA GTGCACTTTTAACCGGACAACTATCAACACGAAGCGGGGAGGAAGCAGGGG
A1.2 293  A              T          C     T     A
A4.1 294  A C   A                   C     AT    A T

A3   339 CTGGAAATGTCCACAGACTTTGCCAAA GACAAAGCCCATAATATCTGAAAGTCAG
A1.2 347                    G        AA TG             T
A4.1 348 T                G  G           TG       G      T

A3   394 TTTCTTC    CATCATTTTGTGTATTAAGGTTCTTTATTCCCCTGTTCTCTGCCTTCCT
A1.2 403 G CT                                  C     T        C
A4.1 403   CT  TCAT                           G C             CA
```

TABLE 2-continued

```
A3   450 GCTTGTCATCTTCACTCATCAGCTGACCATGTTGCCTCTTACGGTGTAAACTTGTACCAG
A1.2 459                                C          GT
A4.1 462                                C  C       T

A3   510 TCTTATGGTCCCTCTGGGCAGTACAGCCATGAATTTGATGGAGACGAGGAGTTCTAT
A1.2 519 T   C          C       C                   T   C       C
A4.1 522     C          C       C                   T   C       C

A3   567 GTGGACCTGGAGAGGAAGGAGACTGTCTGGCAGTTGCCTCTGTTCCGCAGATTTA
A1.2 576                            C     G  G    GA    A   A    G
A4.1 579          G                       TGT      G TC  A ACA

A3   622 GAAGATTTGACCCGCAATTTGCACTGACAAACATCGCTGTGCTAAAACATAACTTGA
A1.2 631   G T            GGG        G       G      GC       C
A4.1 634   ---                                  C

A3   679 ACATCGTGATTAAACGCTCCAACTCTACCGCTGCTACCAATGGTATGTGTCCACCATTCTG
A1.2 688     A            A                                C
A4.1 688    GTC                                                A  A

A3   740 CCTTTCTTTAC     TGATTTATCCCTTTATACCAAGTTTCATTATTTTCTTT
A1.2 749    C       TTAA A GC       CC        G            C
A4.1 749   CC                       C                     A

A3   789 CCAAGAGGTCCCCAGATC 806
A1.2 802                   819
A4.1 798                   815
```

TABLE 3

```
DQB1 Seq

   1 AAGCTTGTGCTCTTTCCATGAATAAATGTCTCTATCTAGGACTCAGAGGT

                     GG              T   T                  A
                                                        G

  51 GTAGG  TCCTTTCCAACATAGAAGGGAGTGA     ACCTCAACGGG ACTTGGGA G
                  TT                                 TT
     C    AC   C   TTT TA C CA AC     GTGA        CA   C
                          A   T                     AT  C          A

 101 GGTAAATCTAGGCATGGGAAGGAAGGTATTTTACCCAGGGACCAAGAGAA

           C
                       G

 151 TACGCGTGTCAGAACGAGGCCAGGCTTAATTCCTGGACCTATCTCGTCAT

     G    A  G   -     A     T                  G
        A               A              T        CG     A

 201 TCCGTTGAACTCTCAGATTTATGTGGATAACTTTATCTCTGAGGTATCCA

       C        G    G                                  C
       C          A  G                 T                T

 251 GGAGCTTCATGAAAAATGGGATTTCATGCGAGAACGCCCTGAT CCCTCTA

         C    G       A
         CA   G                          G          T

 301 AGTGCAGAGGTGCATGTAAAATCAGCCCGACTGCCTCTTCGCTGGGTTCA

             C                                    A  T
             CT                                   C  C

 351 CAGGCTCAGGCAGGGACAGGGCTTTCCTCCCTTTCCTGGATGTAGGAAGG

        CG  A                                    CC
          C                       G              CC  C

 401 C AGATTCCAGAAGCCCGCAAAGAAGGCGGGCAGAGCTGGGCAGAGCCGCC

    CG       C     A  C CG   G          G      -  N N  N
     G       C        C  G   G          G
```

TABLE 3-continued

```
451 GGGAGGATCCCAGGTCTGGAGCGCCAGGCACGGGCGGGCGGGAACTGGAG

                      C      G                          T T
    C       A   A

501 GTCGCGCGGGCGGTTCCACAGCTCCAGGCCGGGTCAGGGCGGCGGCTGCG

              T              G              T
                             G

551 GGGGCGGCCGGGCTGGGGCC             TGACTGACCGGCCGGTGATTCCCCGCAGAG

          A          -     GCA       ---
                         GGGCCGGGGCC

601 GATTTCGTGTACCAGTTTAAGGGCATGTGCTACTTCACCAACGGGACGGA

                                                      A

651 GCGCGTGCGTCTTGTAACCAGACACATCTATAACCGAGAGGAGTACGCGC

                    G G     AG                  A    AT   T
                    G       T                           A

701 GCTTCGACAGCGACGTGGGGGTGTACCGGGCGGTGACGCCGCAGGGGCGG

                           A  T                T  T      T
                              T                          C

751 CCTGTTGCCGAGTACTGGAACAGCCAGAAGGAAGTCCTGGAGAGGACCCG

       CC                                CA             AA
       CC

801 GGCGGAGTTGGA CACGGTGTGCAGACACAACTACGAGGTGGGGTACCGCG

        C G        G                       C  T    A CT     A
                A                          C  T    A CT     A

851 GGATCCTGCAGAGGAGAGGTGAGCTTCGTCGCCCCTCCGTGAGCGC ACCC
                           G
  C  C T      C  C         GG          -T T C    GC C
  C  C T      C  C         G           G         GC C T

901 TTGGCCGGGACCCCGAGTCTCTGTGCCGGGAGGGCG ATGGGGGCGAGGTC

        ------  A         C   A       G  CAA    T  T  C
       A   G   A          CCG         GCGAA        C  C

951 TCTGAAATCTTGAGCCCAGTTCATTCCACCCCAGGGAAAGGAGGCGGCGG
          -
        -C -         C    GG
      G  C          TT                -  CTG C-    A  A

1001    CGGGGGTGGTGGGGGCAGGTGCATCGGAGGGGCGGGGACCTAGGGCAGAG

  CGGT     -  C        T                  A

1051 CAGGGGGACAAGCAGAGTTGGCCAGGCTGCCTAGTGTCCCCCCCAGCCTC

             G           T  A           T  G    - T

1101 CTCGTCCGTCGGCCTCGTCCTCTGCTCTGGACGTTTCTCGCCTCGTGCCT
    C
    C                C            C      -  T

1151 TATGCGTTTGCCTCCTCGTGCCTTACCTTCGCTAAGCAGTTCTCTCTGCC

                                  TA

1201 CCCAGTGCCCACCCTCTTCCCCTGCCCGCCGGCCTCGCTAGCACTGCCCC

       A TT  G                      C   CG             G

1251 ACCCAGCAAGGCCCACAGTCGCGCATTCGCCGCA GGAAGCTT              1292
                        T  CG
       G       T     CTA A AGC CATG AGTGGGAAGCTT
```

TABLE 4

```
DPB1 Seq

DPB4.1 7546                                                GGGAAGATTTGGGAAGAATCGTTAATAT

DPB4.1 7574 TGAGAGAGAGAGGGAGAAAGAGGATTAGATGAGAGTGGCGCCTCCGCTCATGTCCGCCCC

DPB4.1 7634 CTCCCCGCAGAGAATTACCTTTTCCAGGGACGGCAGGAATGCTACGCGTTTAATGGGACA
DPB9        GGAT                 G GCA     TT
New         GGAT                 G GCA     TT
DPw3

BPB4.1 7694 CAGCGCTTCCTGGAGAGATACATCTACAACCGGGAGGAGTTCGCGCGCTTCGACAGCGAC
DPB9                                                       T
New                                                        T
DPw3

DPB4.1 7754 GTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGCTGCGGAGTACTGGAACAGC
DPB9                                                   A  A  C
New                                                   A  A  C
DPw3

DPB4.1 7814 CAGAAGGACATCCTGGAGGAGAAGCGGGCAGTGCCGGACAGGATGTGCAGACACAACTAC
DPB9                              G                    G A
New                 C                                  G A
DPw3                C                                  G A

DPB4.1 7874 GAGCTGGGCGGGCCCATGACCCTGCAGCGCCGAGGTGAGTGAGGGCTTTGGGCCGGCGGT
DPB9              A  A G  G
New               A  A G  G
DPw3              A  A G  G

DPB4.1 7934 CCCAGGGCAGCCCCGCGGGCCCGTGCCCAG
```

Primers for HLA loci

Exemplary HLA locus-specific primers are listed below. Each of the primers hybridizes with at least about 15 consecutive nucleotides of the designated region of the allele sequence. The designation of an exemplary preferred primer together with its sequence is also shown. For many of the primers, the sequence is not identical for all of the other alleles of the locus. For each of the following preferred primers, additional preferred primers have sequences which correspond to the sequences of the homologous region of other alleles of the locus or to their complements.

In one embodiment, Class I loci are amplified by using an A, B or C locus-specific primer together with a Class I locus-specific primer. The Class I primer preferably hybridizes with IVS III sequences (or their complements) or, more preferably, with IVS I sequences (or their complements). The term "Class I-specific primer", as used herein, means that the primer hybridizes with an allele sequence (or its complement) for at least two different Class I loci and does not hybridize with Class II locus allele sequences under the conditions used. Preferably, the Class I primer hybridizes with at least one allele of each of the A, B and C loci. More preferably, the Class I primer hybridizes with a plurality of, most preferably all of, the Class I allele loci or their complements. Exemplary Class I locus-specific primers are also listed below.

HLA Primers

A locus-specific primers

- allelic location: nt 1735–1757 of A3 designation: SGD009.AIVS3.R2NP sequence: CATGTGGCCATCTTGAGAATGGA
- allelic location: nt 1541–1564 of A2 designation: SGD006.AIVS3.R1NP sequence: GCCCGGGAGATCTACAGGCGATCA
- allelic location: nt 1533–1553 of A2 designation: A2.1 sequence: CGCCTCCCTGATCGCCTGTAG
- allelic location: nt 1667–1685 of A2 designation: A2.2 sequence: CCAGAGAGTGACTCTGAGG
- allelic location: nt 1704–1717 of A2 designation: A2.3 sequence: CACAATTAAGGGAT

B locus-specific primers

- allelic location: nt 1108–1131 of B17 designation: SGD007.BIVS3.R1NP sequence: TCCCCGGCGACCTATAGGAGATGG
- allelic location: nt 1582–1604 of B17 designation: SGD010.BIVS3.R2NP sequence: CTAGGACCACCCATGTGACCAGC
- allelic location: nt 500–528 of B27 designation: B2.1 sequence: ATCTCCTCAGACGCCGAGATGCGTCAC
- allelic location: nt 545–566 of B27 designation: B2.2 sequence: CTCCTGCTGCTCTGGGGGGCAG
- allelic location: nt 1852–1876 of B27 designation: B2.3 sequence: ACTTTACCTCCACTCAGATCAGGAG
- allelic location: nt 1945–1976 of B27 designation: B2.4 sequence: CGTCCAGGCTGGTGTCTGGGTTCTGTGCCCCT
- allelic location: nt 2009–2031 of B27 designation: B2.5 sequence: CTGGTCACATGGGTGGTCCTAGG
- allelic location: nt 2054–2079 of B27 designation: B2.6 sequence: CGCCTGAATTTTCTGACTCTTCCCAT

C locus-specific primers

- allelic location: nt 1182–1204 of C3 designation: SGD008.CIVS3.R1NP sequence: ATCCCGGGAGATCTACAGGAGATG
- allelic location: nt 1665–1687 of C3 designation: SGD011.CIVS3.R2NP sequence: AACAGCGCCCATGTGACCATCCT

allelic location: nt 499–525 of C1 designation: C2.1 sequence: CTGGGGAGGCGCCGCGTTGAGGATTCT

allelic location: nt 642–674 of C1 designation: C2.2 sequence: CGTCTCCGCAGTCCCGGTTCTAAAGTTCCCAGT

allelic location: nt 738–755 of C1 designation: C2.3 sequence: ATCCTCGTGCTCTCGGGA

allelic location: nt 1970–1987 of C1 designation: C2.4 sequence: TGTGGTCAGGCTGCTGAC

allelic location: nt 2032–2051 of C1 designation: C2.5 sequence: AAGGTTTGATTCCAGCTT

allelic location: nt 2180–2217 of C1 designation: C2.6 sequence: CCCCTTCCCCACCCCAGGTGTTCCTGTCCATTCTTCAGGA

allelic location: nt 2222–2245 of C1 designation: C2.7 sequence: CACATGGGCGCTGTTGGAGTGTCG

Class I loci-specific primers

allelic location: nt 599–620 of A2 designation: SGD005.IIVS1.LNP sequence: GTGAGTGCGGGGTCGGGAGGGA

allelic location: nt 489–506 of A2 designation: 1.1 sequence: CACCCACCGGGACTCAGA

allelic location: nt 574–595 of A2 designation: 1.2 sequence: TGGCCCTGACCCAGACCTGGGC

allelic location: nt 691–711 of A2 designation: 1.3 sequence: GAGGGTCGGGCGGGTCTCAGC

allelic location: nt 1816–1831 of A2 designation: 1.4 sequence: CTCTCAGGCCTTGTTC

allelic location: nt 1980–1923 of A2 designation: 1.5 sequence: CAGAAGTCGCTGTTCC

DQA1 locus-specific primers

allelic location: nt 23–41 of DQA3 designation: SGD001.DQA1.LNP sequence: TTCTGAGCCAGTCCTGAGA

allelic location: nt 45–64 of DQA3 designation: DQA3 E1a sequence: TTGCCCTGACCACCGTGATG

allelic location: nt 444–463 of DQA3 designation: DQA3 E1b sequence: CTTCCTGCTTGTCATCTTCA

allelic location: nt 536–553 of DQA3 designation: DQA3 E1c sequence: CCATGAATTTGATGGAGA

allelic location: nt 705–723 of DQA3 designation: DQA3 E1d sequence: ACCGCTGCTACCAATGGTA

allelic location: nt 789–806 of DQA3 designation: SGD003.DQA1.RNP sequence: CCAAGAGGTCCCCAGATC

DRA locus-specific primers allelic location: nt 49–68 of DRA HUMMHDRAM (1183 nt sequence, Accession No. K01171) designation: DRA E1 sequence: TCATCATAGCTGTGCTGATG

allelic location: nt 98–118 of DRA HUMMHDRAM (1183 nt sequence, Accession No. K01171) designation: DRA 5'E2 (5' indicates the primer is used as the 5' primer) sequence: AGAACATGTGATCATCCAGGC

allelic location: nt 319–341 of DRA HUMMHDRAM (1183 nt sequence, Accession No. K01171) designation: DRA 3'E2 sequence: CCAACTATACTCCGATCACCAAT

DRB locus-specific primers

allelic location: nt 79–101 of DRB HUMMHDRC (1153 nt sequence, Accession No. K01171) designation: DRB E1 sequence: TGACAGTGACACTGATGGTGCTG

allelic location: nt 123–143 of DRB HUMMHDRC (1153 nt sequence, Accession No. K01171) designation: DRB 5'E2 sequence: GGGGACACCCGACCACGTTTC

allelic location: nt 357–378 of DRB HUMMHDRC (1153 nt sequence, Accession No. K01171) designation: DRB 3'E2 sequence: TGCAGACACAACTACGGGGTTG

DQB1 locus-specific primers

allelic location: nt 509–532 DQB1 DQw1$_v$a designation: DQB E1 sequence: TGGCTGAGGGCAGAGACTCTCCC

allelic location: nt 628–647 of DQB1 DQw1$_v$a designation: DQB 5'E2 sequence: TGCTACTTCACCAACGGGAC

allelic location: nt 816–834 of DQB1 DQw1$_v$a designation: DQB 3'E2 sequence: GGTGTGCACACACAACTAC

allelic location: nt 124–152 of DQB1 DQw1$_v$a designation: DQB 5'IVS1a sequence: AGGTATTTTACCCAGGGACCAAGAGAT

allelic location: nt 314–340 of DQB1 DQw1$_v$a designation: DQB 5'IVS1b sequence: ATGTAAAATCAGCCCGACTGCCTCTTC

allelic location: nt 1140–1166 of DQB1 DQw1$_v$a designation: DQB 3'IVS2 sequence: GCCTCGTGCCTTATGCGTTTGCCTCCT

DPB1 locus-specific primers

allelic location: nt 6116–6136 of DPB14.1 designation: DPB E1 sequence: TGAGGTTAATAAACTGGAGAA

allelic location: nt 7604–7624 of DPB14.1 designation: DPB 5'IVS1 sequence: GAGAGTGGCGCCTCCGCTCAT

allelic location: nt 7910–7929 of DPB14.1 designation: DPB 3'IVS2 sequence: GAGTGAGGGCTTTGGGCCGG

Primer pairs for HLA analyses

It is well understood that for each primer pair, the 5' upstream primer hybridizes with the 5' end of the sequence to be amplified and the 3' downstream primer hybridizes with the complement of the 3' end of the sequence. The primers amplify a sequence between the regions of the DNA to which the primers bind and its complementary sequence including the regions to which the primers bind. Therefore, for each of the primers described above, whether the primer binds to the HLA-encoding strand or its complement depends on whether the primer functions as the 5' upstream primer or the 3' downstream primer for that particular primer pair.

In one embodiment, a Class I locus-specific primer pair includes a Class I locus-specific primer and an A, B or C locus-specific primer. Preferably, the Class I locus-specific primer is the 5' upstream primer and hybridizes with a portion of the complement of IVS I. In that case, the locus-specific primer is preferably the 3' downstream primer and hybridizes with IVS III. The primer pairs amplify a sequence of about 1.0 to about 1.5 Kb.

In another embodiment, the primer pair comprises two locus-specific primers that amplify a DNA sequence that does not include the variable exon(s). In one example of that embodiment, the 3' downstream primer and the 5' upstream primer are Class I locus-specific primers that hybridize with IVS III and its complement, respectively. In that case a sequence of about 0.5 Kb corresponding to the intron sequence is amplified.

Preferably, locus-specific primers for the particular locus, rather than for the HLA class, are used for each primer of the primer pair. Due to differences in the Class II gene sequences, locus-specific primers which are specific for only one locus participate in amplifying the DRB,-DQA1, DQB and DPB loci. Therefore, for each of the preferred Class II locus primer pairs, each primer of the pair participates in amplifying only the designated locus and no other Class II loci.

Analytical methods

In one embodiment, the amplified sequence includes sufficient intron sequences to encompass length polymorphisms. The primer-defined length polymorphisms (PDLPs) are indicative of the HLA locus allele in the sample. For some HLA loci, use of a single primer pair produces primer-defined length polymorphisms that distinguish between some of the alleles of the locus. For other loci, two or more pairs of primers are used in separate amplifications to distinguish the alleles. For other loci, the amplified DNA sequence is cleaved with one or more restriction endonucleases to distinguish the alleles. The primer-defined length polymorphisms are particularly useful in screening processes.

In anther embodiment, the invention provides an improved method that uses PCR amplification of a genomic HLA DNA sequence of one HLA locus. Following amplification, the amplified DNA sequence is combined with at least one endonuclease to produce a digest. The endonuclease cleaves the amplified DNA sequence to yield a set of fragments having distinctive fragment lengths. Usually the amplified sequence is divided, and two or more endonuclease digests are produced. The digests can be used, either separately or combined, to produce RFLP patterns that can distinguish between individuals. Additional digests can be prepared to provide enhanced specificity to distinguish between even closely related individuals with the same HLA type.

In a preferred embodiment, the presence of a particular allele can be verified by performing a two step amplification procedure in which an amplified sequence produced by a first primer pair is amplified by a second primer pair which binds to and defines a sequence within the first amplified sequence. The first primer pair can be specific for one or more alleles of the HLA locus. The second primer pair is preferably specific for one allele of the HLA locus, rather than a plurality of alleles. The presence of an amplified sequence indicates the presence of the allele, which is confirmed by production of characteristic RFLP patterns.

To analyze RFLP patterns, fragments in the digest are separated by size and then visualized. In the case of typing for a particular HLA locus, the analysis is directed to detecting the two DNA allele sequences that uniquely characterize that locus in each individual. Usually this is performed by comparing the sample digest RFLP patterns to a pattern produced by a control sample of known HLA allele type. However, when the method is used for paternity testing or forensics, the analysis need not involve identifying a particular locus or loci but can be done by comparing single or multiple RFLP patterns of one individual with that of another individual using the same restriction endonuclease and primers to determine similarities and differences between the patterns.

The number of digests that need to be prepared for any particular analysis will depend on the desired information and the particular sample to be analyzed. For example, one digest may be sufficient to determine that an individual cannot be the person whose blood was found at a crime scene. In general, the use of two to three digests for each of two to three HLA loci will be sufficient for matching applications (forensics, paternity). For complete HLA haplotyping; e.g., for transplantation, additional loci may need to be analyzed.

As described previously, combinations of primer pairs can be used in the amplification method to amplify a particular HLA DNA locus irrespective of the allele present in the sample. In a preferred embodiment, samples of HLA DNA are divided into aliquots containing similar amounts of DNA per aliquot and are amplified with primer pairs (or combinations of primer pairs) to produce amplified DNA sequences for additional HLA loci. Each amplification mixture contains only primer pairs for one HLA locus. The amplified sequences are preferably processed concurrently, so that a number of digest RFLP fragment patterns can be produced from one sample. In this way, the HLA type for a number of alleles can be determined simultaneously.

Alternatively, preparation of a number of RFLP fragment patterns provides additional comparisons of patterns to distinguish samples for forensic and paternity analyses where analysis of one locus frequently fails to provide sufficient information for the determination when the sample DNA has the same allele as the DNA to which it is compared.

The use of HLA types in paternity tests or transplantation testing and in disease diagnosis and prognosis is described in Basic & Clinical Immunology, 3rd Ed (1980) Lange Medical Publications, pp 187–190, which is incorporated herein by reference in its entirety. HLA determinations fall into two general categories. The first involves matching of DNA from an individual and a sample. This category involves forensic determinations and paternity testing. For category 1 analysis, the particular HLA type is not as important as whether the DNA from the individuals is related. The second category is in tissue typing such as for use in transplantation. In this case, rejection of the donated blood or tissue will depend on whether the recipient and the donor express the same or different antigens. This is in contrast to first category analyses where differences in the HLA DNA in either the introns or exons is determinative.

For forensic applications, analysis of the sample DNA of the suspected perpetrator of the crime and DNA found at the crime scene are analyzed concurrently and compared to determine whether the DNA is from the same individual. The determination preferably includes analysis of at least three digests of amplified DNA of the DQA1 locus and preferably also of the A locus. More preferably, the determination also includes analysis of at least three digests of amplified DNA of an additional locus, e.g. the DPB locus. In this way, the probability that differences between the DNA samples can be discriminated is sufficient.

For paternity testing, the analysis involves comparison of DNA of the child, the mother and the putative father to determine the probability that the child inherited the obligate haplotype DNA from the putative father. That is, any DNA sequence in the child that is not present in the mother's DNA must be consistent with being provided by the putative father. Analysis of two to three digests for the DQA1 and preferably also for the A locus is usually sufficient. More preferably, the determination also includes analysis of digests of an additional locus, e.g. the DPB locus.

For tissue typing determinations for transplantation matching, analysis of three loci (HLA A, B, and DR) is often

sufficient-. Preferably, the final analysis involves comparison of additional loci including DQ and DP.

Production of RFLP fragment patterns

The following table of exemplary fragment pattern lengths demonstrates distinctive patterns. For example, as shown in the table, BsrI cleaves A2, A3 and A9 allele amplified sequences defined by primers SGD005.IIVS1.LNP and SGD009.AIVS3.R2NP into sets of fragments with the following numbers of nucleotides (740, 691), (809, 335, 283) and (619, 462, 256, 93), respectively. The fragment patterns clearly indicate which of the three A alleles is present. The following table illustrates a number of exemplary endonucleases that produce distinctive RFLP fragment patterns for exemplary A allele sequences.

Table 2 illustrates the set of RFLP fragments produced by use of the designated endonucleases for analysis of three A locus alleles. For each endonuclease, the number of nucleotides of each of the fragments in a set produced by the endonuclease is listed. The first portion of the table illustrates RFLP fragment lengths using the primers designated SGD009.AIVS3.R2NP and SGD005.IIVS1.LNP which produce the longer of the two exemplary sequences. The second portion of the table illustrates RFLP fragment lengths using the primers designated SGD006.AIVS3.R1NP and SGD005.IIVS1.LNP which produce the shorter of the sequences. The third portion of the table illustrates the lengths of fragments of a DQA1 locus-specific amplified sequence defined by the primers designated SGD001.DQA1.LNP and SGD003.DQA1.RNP.

As shown in the Table, each of the endonucleases produces a characteristic RFLP fragment pattern which can readily distinguish which of the three A alleles is present in a sample.

TABLE 5

RFLP FRAGMENT PATTERNS

					A — Long						
BsrI	A2		740	691							
	A3	809					335	283			
	A9				619	462			256		93
Cfr101	A2	1055			399	245					
	A3			473	399	247					
	A9		786		399						
DraII	A2	698					251		138		
	A3				369	315	251	247			
	A9		596	427			251			80	
FokI	A2		728		248			151			
	A3			515		225	213	151			
	A9	1004						151			
GsuI	A2		868		547				36		
	A3	904				523					
	A9			638			419	373			
HphI	A2	1040					239			72	
	A3			419	375			218	163		
	A9		643	419	373						
MboII	A2		1011			165	143	132			
	A3			893	194		143		115		
	A9	1349								51	
PpumI	A2	698					295	251		138	
	A3				369	364		251	242		
	A9		676	503				251			
PssI	A2	695					295	251		138	
	A3				366	315		251	242		
	A9		596	427				251			
					A — Short						
BsrI	A2	691					254				
	A3			345	335	283					
	A9		619				256	93			
Cfr101	A2										
	A3										
	A9										
DraII	A2			295	251	210	138				
	A3		315		251	210					
	A9	427			251	210					
FokI	A2		293	248			151	143	129	51	
	A3				225	213	151	143	129	51	
	A9	539					151	146	129		
GsuI	A2		868				61	36			
	A3	904					59				
	A9			414	373	178					
HphI	A2	554			339						
	A3		411	375		177					
	A9		414	373		178					
MboII	A2										
	A3										
	A9										
PpumI	A2			295	257		212		69		

TABLE 5-continued

	RFLP FRAGMENT PATTERNS																				
	A3			364					251		210		72		66						
	A9	503							251		211										
PssI	A2					295		251	219				72								
	A3			315				251			207		72		66						
	A9	427						251			208		72								
								DQA1													
AluI																					
DQA3		449		335																	
DQA4.1			338	332					122												
DQA1.2				335	287				123						52						
CvijI																					
DQA3						271		187	122		99			64		34				7	
DQA4.1						277	219			102		79			55	36		17		7	
DQA1.2								201		101	99	80	76		55	36	35			7	
DdeI																					
DQA3	587											88		65			30		11	3	
DQA4.1			388					194				89		64		36			11	3	
DQA1.2			395					165				88		65	41	36			11	3	
MboII																					
DQA3			366					184	172					62							
DQA4.1		407	353													32					
DQA1.2			330	316								89				32	30				
MnlI																					
DQA3							214		176	172			72		43	36	23	21	17	10	
DQA4.1					294				179		149				40	36	33	21			
DQA1.2							216			136	123	73	54	44	40	36	24	21	15	10	5
NlaIII																					
DQA3		458				266							60								
DQA4.1					300	263	229														
DQA1.2							223	190			124	116	75			39	30				
TthIIII																					
DQA3		417					226			141											
DQA4.1	426		371																		
DQA1.2	428								148	141			75								

Screening Analysis for Genetic Disease

Carriers of genetic diseases and those affected by the disease can be identified by use of the present method. Depending on the disease, the screening analysis can be used to detect the presence of one or more alleles associated with the disease or the presence of haplotypes associated with the disease. Furthermore, by analyzing haplotypes, the method can detect genetic diseases that are not associated with coding region variations but are found in regulatory or other untranslated regions of the genetic locus. The screening method is exemplified below by analysis of cystic fibrosis (CF).

Cystic fibrosis is an autosomal recessive disease, requiring the presence of a mutant gene on each chromosome. CF is the most common genetic disease in Caucasians, occurring once in 2,000 live births. It is estimated that one in forty Caucasians are carriers for the disease.

Recently a specific deletion of three adjacent basepairs in the open reading frame of the putative CF gene leading to the loss of a phenylalanine residue at position 508 of the predicted 1480 amino acid polypeptide was reported [Kerem et al, *Science* 245:1073–1080 (1989)]. Based on haplotype analysis, the deletion may account for most CF mutations in Northern European populations (about 68%). A second mutation is reportedly prevalent in some Southern European populations. Additional data indicate that several other mutations may cause the disease.

Studies of haplotypes of parents of CF patients (who necessarily have one normal and one disease-associated haplotype) indicated that there are at least 178 haplotypes associated with the CF locus. Of those haplotypes, 90 are associated only with the disease; 78 are found only in normals; and 10 are associated with both the disease and with normals (Kerem et al, supra). The disease apparently is caused by several different mutations, some in very low frequency in the population. As demonstrated by the haplotype information, there are more haplotypes associated with the locus than there are mutant alleles responsible for the disease.

A genetic screening program (based on amplification of exon regions and analysis of the resultant amplified DNA sequence with probes specific for each of the mutations or with enzymes producing RFLP patterns characteristic of each mutation) may take years to develop. Such tests would depend on detection and characterization of each of the mutations, or at least of mutations causing about 90 to 95% or more of the cases of the disease. The alternative is to detect only 70 to 80% of the CF-associated genes. That alternative is generally considered unacceptable and is the cause of much concern in the scientific community.

The present method directly determines haplotypes associated with the locus and can detect haplotypes among the

178 currently recognized haplotypes associated with the disease locus. Additional haplotypes associated with the disease are readily determined through the rapid analysis of DNA of numerous CF patients by the methods of this invention. Furthermore, any mutations which may be associated with noncoding regulatory regions can also be detected by the method and will be identified by the screening process.

Rather than attempting to determine and then detect each defect in a coding region that causes the disease, the present method amplifies intron sequences associated with the locus to determine allelic and sub-allelic patterns. In contrast to use of mutation-specific probes where only known sequence defects can be detected, new PDLP and RFLP patterns produced by intron sequences indicate the presence of a previously unrecognized haplotype.

The same analysis can be performed for phenylalanine hydroxylase locus nutations that cause phenylketonuria and for beta-globin mutations that cause beta-thalassemia and sickle cell disease and for other loci known to be associated with a genetic disease. Furthermore, neither the mutation site nor the location for a disease gene is required to determine haplotypes associated with the disease. Amplified intron sequences in the regions of closely flanking RFLP markers, such as are known for Huntington's disease and many other inherited diseases, can provide sufficient information to screen for haplotypes associated with the disease.

Muscular dystrophy (MD) is a sex-linked disease. The disease-associated gene comprises a 2.3 million basepair sequence that encodes 3,685 amino acid protein, dystrophin. A map of mutations for 128 of 34 patients with Becker's muscular dystrophy and 160 patients with Duchenne muscular dystrophy identified 115 deletions and 13 duplications in the coding region sequence [Den Dunnen et al, *Am. J. Hum. Genet.* 45:835–847 (1989)]. Although the disease is associated with a large number of mutations that vary widely, the mutations have a non-random distribution in the sequence and are localized to two major mutation hot spots, Den Dunnen et al, supra. Further, a recombination hot spot within the gene sequence has been identified [Grimm et al, *Am. J. Hum. Genet.* 45:368–372 (1989)].

For analysis of MD, haplotypes on each side of the recombination hot spot are preferably determined. Primer pairs defining amplified DNA sequences are preferably located near, within about 1 to 10 Kbp of the hot spot on either side of the hot spot. In addition, due to the large size of the gene, primer pairs defining amplified DNA sequences are preferably located near each end of the gene sequence and most preferably also in an intermediate location on each side of the hot spot. In this way, haplotypes associated with the disease can be identified.

Other diseases, particularly malignancies, have been shown to be the result of an inherited recessive gene together with a somatic mutation of the normal gene. One malignancy that is due to such "loss of heterogeneity" is retinoblastoma, a childhood cancer. The loss of the normal gene through mutation has been demonstrated by detection of the presence of one mutation in all somatic cells (indicating germ cell origin) and detection of a second mutation in some somatic cells [Scheffer et al, *Am. J. Hum. Genet.* 45:252–260 (1989)]. The disease can be detected by amplifying somatic cell, genomic DNA sequences that encompass sufficient intron sequence nucleotides. The amplified DNA sequences preferably encompass intron sequences locate near one or more of the markers described by Scheffer et al, supra. Preferably, an amplified DNA sequence located near an intragenic marker and an amplified DNA sequence located near a flanking marker are used.

An exemplary analysis for CF is described in detail in the examples. Analysis of genetic loci for other monogenic and multigenic genetic diseases can be performed in a similar manner.

As the foregoing description indicates, the present method of analysis of intron sequences is generally applicable to detection of any type of genetic trait. Other monogenic and multigenic traits can be readily analyzed by the methods of the present invention. Furthermore, the analysis methods of the present method are applicable to all eukaryotic cells, and are preferably used on those of plants and animals. Examples of analysis of BoLA (bovine MHC determinants) further demonstrates the general applicability of the methods of this invention.

This invention is further illustrated by the following specific but non-limiting examples. Procedures that are constructively reduced to practice are described in the present tense, and procedures that have been carried out in the laboratory are set forth in the past tense.

EXAMPLE 1

Forensic Testing

DNA extracted from peripheral blood of the suspected perpetrator of a crime and DNA from blood found at the crime scene are analyzed to determine whether the two samples of DNA are from the same individual or from different individuals.

The extracted DNA from each sample is used to form two replicate aliquots per sample, each aliquot having 1 μg of sample DNA. Each replicate is combined in a total volume of 100 μl with a primer pair (1 μg of each primer), dNTPs (2.5 mM each) and 2.5 units of Taq polymerase in amplification buffer (50 mM KCl; 10 mM Tris-HCl, pH 8.0; 2.5 mM $MgCl_2$; 100 μg/ml gelatin) to form four amplification reaction mixtures. The first primer pair contains the primers designated SGD005.IIVS1.LNP and SGD009.AIVS3.R2NP (A locus-specific). The second primer pair contains the primers designated SGD001.DQA1.LNP and SGD003.DQA1.RNP (DQA locus-specific). Each primer is synthesized using an Applied Biosystems model 308A DNA synthesizer. The amplification reaction mixtures are designated SA (suspect's DNA, A locus-specific primers), SD (suspect's DNA, DQA1 locus-specific primers), CA (crime scene DNA, A locus-specific primers) and CD (crime scene DNA, DQA1 locus-specific primers).

Each amplification reaction mixture is heated to 94° C. for 30 seconds. The primers are annealed to the sample DNA by cooling the reaction mixtures to 65° C. for each of the A locus-specific amplification mixtures and to 55° C. for each of the DQA1 locus-specific amplification mixtures and maintaining the respective temperatures for one minute. The primer extension step is performed by heating each of the amplification mixtures to 72° C. for one minute. The denaturation, annealing and extension cycle is repeated 30 times for each amplification mixture.

Each amplification mixture is aliquoted to prepare three restriction endonuclease digestion mixtures per amplification mixture. The A locus reaction mixtures are combined with the endonucleases BsrI, Cfr101 and DraII. The DQA1 reaction mixtures are combined with AluI, CvijI and DdeI.

To produce each digestion mixture, each of three replicate aliquots of 10 μl of each amplification mixture is combined with 5 units of the respective enzyme for 60 minutes at 37° C. under conditions recommended by the manufacturer of each endonuclease.

Following digestion, the three digestion mixtures for each of the samples (SA, SD, CA and CD) are pooled and electrophoresed on a 6.5% polyacrylamide gel for 45 minutes at 100 volts. Following electrophoresis, the gel is stained with ethidium bromide.

The samples contain fragments of the following lengths:

SA: 786, 619, 596, 462, 427, 399, 256, 251, 93, 80 CA: 809, 786, 619, 596, 473, 462, 427, 399, 369, 335, 315, 283, 256, 251, 247, 93, 80

SD: 388, 338, 332, 277, 219, 194, 122, 102, 89, 79, 64, 55 CD: 587, 449, 388, 338, 335, 332, 277, 271, 219, 194, 187, 122, 102, 99, 89, 88, 79, 65, 64, 55

The analysis demonstrates that the blood from the crime scene and from the suspected perpetrator are not from the same individual. The blood from the crime scene and from the suspected perpetrator are, respectively, A3, A9, DQA1 0501, DQA1 0301 and A9, A9, DQA1 0501, DQA1 0501.

EXAMPLE 2

Paternity Testing

Chorionic villus tissue was obtained by transcervical biopsy from a 7-week old conceptus (fetus). Blood samples were obtained by venepuncture from the mother (M), and from the alleged father (AF). DNA was extracted from the chorionic villus biopsy, and from the blood samples. DNA was extracted from the sample from M by use of nonionic detergent (Tween 20) and proteinase K. DNA was extracted from the sample from F by hypotonic lyuis. More specifically, 100 μl of blood was diluted to 1.5 ml in PBS and centrifuged to remove buffy coat. Following two hypotonic lysis treatments involving resuspension of buffy coat cells in water, the pellets were washed until redness disappeared. Colorless pellets were resuspended in water and boiled for 20 minutes. Five 10 mm chorionic villus fronds were received. One frond was immersed in 200 μl water. NaOH was added to 0.05 M. The sample was boiled for 20 minutes and then neutralized with HCl. No further purification was performed for any of the samples.

The extracted DNA was submitted to PCR for amplification of sequences associated with the HLA loci, DQA1 and DPB1. The primers used were: (1) as a 5' primer for the DQA1 locus, the primer designated SGD001.DQA1.LNP (DQA 5'IVS1) (corresponding to nt 23–39 of the DQA1 0301 allele sequence) and as the 3' primer for the DQA1 locus, the primer designated SGD003.DQA1.RNP (DQA 3'IVS2 corresponding to nt 789–806 of the DQA1 0301 sequence; (2) as the DPB primers, the primers designated 5'IVS1 nt 7604–7624 and 3'IVS2 7910–7929. The amplification reaction mixtures were: 150 ng of each primer; 25 μ of test DNA; 10 mM Tris HCl, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$; 0.01% (w/v) gelatin; 200 μM dNTPs; water to 100 μl and 2.5 U Taq polymerase.

The amplification was performed by heating the amplification reaction mixture to 94° C. for 10 minutes prior to addition of Taq polymerase. For DQA1, the amplification was performed at 94° C. for 30 seconds, then 55° C. for 30 seconds, then 72° C. for 1 minute for 30 cycles, finishing with 72° C. for 10 minutes. For DPB, the amplification was performed at 96° C. for 30 seconds, then 65° C. for 30 seconds, finishing with 65° C. for 10 minutes.

Amplification was shown to be technically satisfactory by test gel electrophoresis which demonstrated the presence of double stranded DNA of the anticipated size in the amplification reaction mixture. The test gel was 2% agarose in TBE (tris borate EDTA) buffer, loaded with 15 μl of the amplification reaction mixture per lane and electrophoresed at 200 v for about 2 hours until the tracker dye migrated between 6 to 7 cm into the 10 cm gel.

The amplified DQA1 and DPB1 sequences were subjected to restriction endonuclease digestion using DdeI and MboII (8 and 12 units, respectively at 37° C. for 3 hours) for DQA1, and RsaI and FokI (8 and 11 units, respectively at 37° C. overnight) for DPB1 in 0.5 to 2.0 μl of enzyme buffers recommended by the supplier, Pharmacia together with 16–18 μl of the amplified product. The digested DNA was fragment size-length separated on gel electrophoresis (3% Nusieve). The RFLP patterns were examined under ultraviolet light after staining the gel with ethidium bromide.

Fragment pattern analysis is performed by allele assignment of the non-maternal alleles using expected fragment sizes based on the sequences of known endonuclease restriction sites. The fragment pattern analysis revealed the obligate paternal DQA1 allele to be DQA1 0102 and DPB to be DPwl. The fragment patterns were consistent with AF being the biological father.

To calculate the probability of true paternity, HLA types were assigned. Maternal and AF DQA1 types were consistent with those predicted from the HLA Class II gene types determined by serological testing using lymphocytotoxic antisera.

Considering alleles of the two HLA loci as being in linkage equilibrium, the combined probability of non-paternity was given by:

$$0.042 \times 0.314 - 0.013$$

i.e. the probability of paternity is (1–0.013) or 98.7%.

The relative chance of paternity is thus 74:75, i.e. the chance that the AF is not the biological father is approximately 1 in 75. The parties to the dispute chose to regard these results as confirming the paternity of the fetus by the alleged father.

EXAMPLE 3

Analysis of the HLA DQA1 Locus

The three haplotypes of the HLA DQA1 0102 locus were analyzed as described below. Those haplotypes are DQA1 0102 DR15 Dw2; DQA1 0102 DR16 Dw21; and DQA1 0102 DR13 Dw19. The distinction between the haplotypes is particularly difficult because there is a one basepair difference between the 0102 alleles and the 0101 and 0103 alleles, which difference is not unique in DQA1 allele sequences.

The procedure used for the amplification is the same as that described in Example 1, except that the amplification used thirty cycles of 94° C. for 30 seconds, 60° C. for 30 seconds, and 72° C. for 60 seconds. The sequences of the primers were:

SGD 001 — 5' TTCTGAGCCAGTCCTGAGA 3'; and
SGD 003 — 5' GATCTGGGGACCTCTTGG 3'.

These primers hybridize to sequences about 500 bp upstream from the 5' end of the second exon and 50 bp downstream from the second exon and produce amplified DNA sequences in the 700 to 800 bp range.

Following amplification, the amplified DNA sequences were electrophoresed on a 4% polyacrylamide gel to determine the PDLP type. In this case, amplified DNA sequences for 0102 comigrate with (are the same length as) 0101 alleles and subsequent enzyme digestion is necessary to distinguish them.

The amplified DNA sequences were digested using the restriction enzyme AluI (Bethesda Research Laboratories) which cleaves DNA at the sequence AGCT. The digestion was performed by mixing 5 units (1 μl) of enzyme with 10 μl of the amplified DNA sequence (between about 0.5 and 1 μg of DNA) in the enzyme buffer provided by the manufacturer according to the manufacturer's directions to form a digest. The digest was then incubated for 2 hours at 37° C. for complete enzymatic digestion.

The products of the digestion reaction are mixed with approximately 0.1 μg of "ladder" nucleotide sequences (nucleotide control sequences beginning at 123 bp in length and increasing in length by 123 bp to a final size of about 5,000 bp; available commercially from Bethesda Research Laboratories, Bethesda Md.) and were electrophoresed using a 4% horizontal ultra-thin polyacrylamide gel (E-C Apparatus, Clearwater Fla.). The bands in the gel were visualized (stained) using silver stain technique [Allen et al, *BioTechniques* 7:736–744 (1989)].

Three distinctive fragment patterns which correspond to the three haplotypes were produced using AluI. The patterns (in base pair sized fragments) were:

1. DR15 DQ6 Dw2: 120, 350, 370, 480
2. DR13 DQ6 Dw19: 120, 330, 350, 480
3. DR16 DQ6 Dw21: 120, 330, 350

The procedure was repeated using a 6.5% vertical polyacrylamide gel and ethidium bromide stain and provided the same results. However, the fragment patterns were more readily distinguishable using the ultrathin gels and silver stain.

This exemplifies analysis according to the method of this invention. Using the same procedure, 20 of the other 32 DR/DQ haplotypes for DQA1 were identified using the same primer pair and two additional enzymes (DdeI and MboII). PDLP groups and fragment patterns for each of the DQA1 haplotypes with the three endonucleases are illustrated in Table 6.

AluI

PDLP	DR	Dw	480	410	405	400	390	370	360	350	340	330	310	300	270	240	120	110	100
0101	1	1	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	+	−
	1	20	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	+	−
	14	9	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	−	−
0102	15	2	+	−	−	−	−	+	−	+	−	−	−	−	−	−	+	−	−
	16	21	−	−	−	−	−	−	−	+	−	+	−	−	−	−	+	−	−
	13	19	+	−	−	−	−	−	−	+	−	+	−	−	−	−	+	−	−
0103	13	18.24	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	13	18.25	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	+	−
	8	8.3	−	−	−	−	−	−	+	−	+	−	−	−	−	−	+	−	+
	11	DB2	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	15	12	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0201	7	DB1	−	−	−	+	−	−	−	−	−	−	−	+	−	−	−	−	−
	7	17	−	−	−	+	−	−	−	−	−	−	−	+	−	−	−	−	−
	7	11	−	−	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−
0301	4	4(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	13.2(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	4(8)	−	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−	−
	4	10(8)	−	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−	−
	4	13.1(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	14(8)	−	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−	−
	4	KT2	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	15	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	9	23	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−	−	−
0401	8	8.1	−	(+)	−	−	−	−	−	−	−	−	+	−	−	+	+	+	−
	8	8.2	−	(+)	−	−	−	−	−	−	−	−	+	−	−	+	+	+	−
	3	RSH	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0501	3	3,24	−	(+)	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	3	3,25	−	(+)	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	11	5	−	(+)	+	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	11	5(9104)	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	14	16	−	(+)	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	12	DB6	−	(+)	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−
	16	22	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0601	8	8.3	−	(+)	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−

DdeI

PDLP	DR	Dw	650	520	450	420	410	400	390	340	300	200	190	150	90	80	60	50
0101	1	1	−	−	−	+	−	−	−	−	−	−	−	+	−	+	+	−
	1	20	−	−	−	+	−	−	−	−	−	−	−	+	−	+	+	−
	14	9	−	−	−	+	−	−	−	−	−	−	−	+	−	+	+	−
0102	15	2	−	+	−	+	−	−	−	−	−	−	−	+	−	+	+	−
	16	21	−	−	−	−	+	−	−	−	−	−	−	+	−	+	+	−
	13	19	−	−	−	−	−	−	−	−	−	−	−	+	−	+	+	−

-continued

MboII

PDLP	DR	Dw	390	385	380	370	365	360	350	340	335	330	305	300	250	190	180	170	140	130
0101	1	1	−	−	−	−	−	−	+	−	−	−	−	+	−	−	−	−	−	−
	1	20	−	−	−	−	−	−	+	−	−	−	−	+	−	−	−	−	−	−
	14	9	−	−	−	+	−	−	+	−	−	−	+	−	−	−	−	−	−	−
0102	15	2	−	−	−	+	−	−	−	−	−	−	−	+	−	−	−	−	−	−
	16	21	−	−	−	+	−	−	−	−	−	−	−	+	−	−	−	−	−	−
	13	19	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0103	13	18,24	−	−	−	−	−	+	−	−	−	−	−	−	−	+	−	−	+	−
	13	18,25	−	−	−	−	−	+	−	−	−	−	−	−	−	+	−	−	+	−
	8	8.3	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	11	DB2	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	15	12	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0201	7	DB1	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−	−
	7	17	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−	−
	7	11	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−	−
0301	4	4(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	13.2(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	4(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	10(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	13.1(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	14(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	KT2	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	15	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	9	23	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0401	8	8.1	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	8	8.2	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	3	RSH	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0501	3	3,24	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	3	3,25	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	11	5	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	11	5(9104)	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−
	14	16	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	12	DB6	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	16	22	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0601	8	8.3	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−

PDLP	DR	Dw	650	520	450	420	410	400	390	300	200	150	90	80	60	50
0103	13	18,24	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	13	18,25	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	8	8.3	−	−	+	−	−	−	−	−	−	+	−	+	−	−
	11	DB2	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	15	12	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0201	7	DB1	−	−	−	−	−	+	−	−	−	+	−	+	−	−
	7	17	−	−	−	−	−	−	−	−	−	+	−	+	−	−
	7	11	−	−	−	−	−	+	−	−	−	+	−	+	−	−
0301	4	4(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	13.2(7)	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	4(8)	−	−	−	−	−	−	−	−	−	−	+	−	−	+
	4	10(8)	+	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	13.1(8)	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	14(8)	+	−	−	−	−	−	−	−	−	+	−	+	−	−
	4	KT2	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	4	15	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	9	23	+	−	−	−	−	−	−	−	+	−	+	−	−	−
0401	8	8.1	−	−	−	−	−	−	+	+	+	+	+	+	−	−
	8	8.2	−	−	−	−	−	−	+	+	−	+	−	+	−	−
	3	RSH	−	−	−	−	−	−	−	−	−	−	−	−	−	−
0501	3	3,24	−	−	−	−	−	+	−	−	+	−	−	+	−	−
	3	3,25	−	−	−	−	−	+	−	−	+	−	−	+	−	−
	11	5	−	−	−	−	−	+	−	−	+	−	−	+	−	−
	11	5(9104)	−	−	−	−	−	+	−	−	+	−	−	+	−	−
	14	16	−	−	−	−	−	−	+	−	−	−	−	−	−	−
	12	DB6	−	−	−	−	−	+	−	−	+	−	−	+	−	−
	16	22	−	−	−	−	−	+	−	−	+	−	−	+	−	−
0601	8	8.3	−	−	−	−	−	−	−	+	+	−	+	−	−	+

-continued

PDLP	DR	Dw	390	385	380	370	365	360	350	340	335	330	305	300	250	190	180	170	140	130
	4	13.1(8)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	4	14(8)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+
	4	KT2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	4	15	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	9	23	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
0401	8	8.1	–	–	–	–	–	+	–	–	–	+	–	–	+	–	+	+	–	–
	8	8.2	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–
	3	RSH	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–
0501	3	3.24	+	–	+	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–
	3	3.25	+	–	+	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–
	11	5	+	–	+	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–
	11	5(9104)	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	+	–
	14	16	–	+	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–
	12	DB6	+	–	+	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–
	16	22	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
0601	8	8.3	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–

This example illustrates the ability of the method of this invention to distinguish the alleles and haplotypes of a genetic locus. Specifically, the example shows that PDLP analysis stratifies five of the eight alleles. These three restriction endonuclease digests distinguish each of the eight alleles and many of the 35 known haplotypes of the locus. The use of additional endonuclease digests for this amplified DNA sequence can be expected to distinguish all of the known haplotypes and to potentially identify other previously unrecognized haplotypes. Alternatively, use of the same or other endonuclease digests for another amplified DNA sequence in this locus can be expected to distinguish the haplotypes.

In addition, analysis of amplified DNA sequences at the DRA locus in the telomeric direction and DQB in the centromeric direction, preferably together with analysis of a central locus, can readily distinguish all of the haplotypes for the region.

The same methods are readily applied to other loci.

EXAMPLE 4

Analysis of the HLA DQA1 Locus

The DNA of an individual is analyzed to determine which of the three haplotypes of the HLA DQA1 0102 locus are present. Genomic DNA is amplified as described in Example 3. Each of the amplified DNA sequences is sequenced to identify the haplotypes of the individual. The individual is shown to have the haplotypes DR15 DQ6 Dw2; DR13 DQ6 Dw19.

The procedure is repeated as described in Example 3 through the production of the AluI digest. Each of the digest fragments is sequenced. The individual is shown to have the haplotypes DR15 DQ6 Dw2; DR13 DQ6 Dw19.

EXAMPLE 5

DQA1 Allele-Specific Amplification

Primers were synthesized that specifically bind the 0102 and 0301 alleles of the DQA1 locus. The 5' primer was the SGD 001 primer used in Example 3. The sequences of the 3' primers are listed below.

0102 5' TTGCTGAACTCAGGCCACC 3'
0301 5' TGCGGAACAGAGGCAACTG 3'

The amplification was performed as described in Example 3 using 30 cycles of a standard (94° C., 60° C., 72° C.) PCR reaction. The template DNAs for each of the 0101, 0301 and 0501 alleles were amplified separately. As determined by gel electrophoresis, the 0102-allele-specific primer amplified only template 0102 DNA and the 0301-allele-specific primer amplified only template 0301 DNA. Thus, each of the primers was allele-specific.

EXAMPLE 6

Detection of Cystic Fibrosis

The procedure used for the amplification described in Example 3 is-repeated. The sequences of the primers are illustrated below. The first two primers are upstream primers, and the third is a downstream primer. The primers amplify a DNA sequence that encompasses all of intervening sequence 1

5' CAG AGG TCG CCT CTG GA 3';
5' AAG GCC AGC GTT GTC TCC A 3'; and
3' CCT CAA AAT TGG TCT GGT 5'.

These primers hybridize to the complement of sequences located from nt 136–152 and nt 154–172, and to nt 187–207. [The nucleotide numbers are found in Riordan et al, *Science* 245:1066–1072 (1989).]

Following amplification, the amplified DNA sequences are electrophoresed on a 4% polyacrylamide gel to determine the PDLP type. The amplified DNA sequences are separately digested using each of the restriction enzymes AluI, MnlI and RsaI (Bethesda Research Laboratories). The digestion is performed as described in Example 3. The products of the digestion reaction are electrophoresed and visualized using a 4% horizontal ultra-thin polyacrylamide gel and silver stain as described in Example 3.

Distinctive fragment patterns which correspond to disease-associated and normal haplotypes are produced.

EXAMPLE 7

Analysis of Bovine Leukocyte Antigen Class I

Bovine Leukocyte Antigen (BOLA) Class I alleles and haplotypes are analyzed in the same manner as described in Example 3. The primers are listed below.

Bovine Primers (Class I HLA homolog)		T_m
5' primer: 5' TCC TGG TCC TGA CCG AGA 3'		(62°)
3' primer: 1)	3' A TGT GCC TTT GGA GGG TCT 5' (for ~600 bp product)	(62°)
2)	3' GCC AAC AT GAT CCG CAT 5' (for ~900 bp product)	(62°)

For the approximately 900 bp sequence PDLP analysis is sufficient to distinguish alleles 1 and 3 (893 and 911 bp, respectively). Digests are prepared as described in Example 3 using AluI and DdeI. The following patterns are produced for the 900 bp sequence.

Allele 1, AluI digest: 712, 181 Allele 3, AluI digest: 430, 300, 181

Allele 1, DdeI digest: 445, 201, 182, 28 Allele 3, DdeI digest: 406, 185, 182, 28, 16

The 600 bp sequence also produces distinguishable fragment patterns for those alleles. However, those patterns are not as dramatically different as the patterns produced by the 600 bp sequence digests.

EXAMPLE 8

Preparation of Primers

Each of the following primers is synthesized using an Applied Biosystems model 308A DNA synthesizer.

HLA locus primers

A locus-specific primers

SGD009.AIVS3.R2NP CATGTGGCCATCTTGAGAATGGA SGD006.AIVS3.R1NP GCCCGGGAGATCTACAGGCGATCA A2.1 CGCCTCCCTGATCGCCTGTAG A2.2

CCAGAGAGTGACTCTGAGG A2.3 CACAATTAAGGGAT

B locus-specific primers

SGD007.BIVS3.R1NP TCCCCGGCGACCTATAGGAGATGG SGD010.BIVS3.R2NP CTAGGACCACCCATGTGACCAGC B2.1 ATCTCCTCAGACGCCGAGATGCGTCAC B2.2 CTCCTGCTGCTCTGGGGGGCAG B2.3 ACTTTACCTCCACTCAGATCAGGAG B2.4 CGTCCAGGCTGGTGTCTGGGTTCTGTGCCCCT B2.5 CTGGTCACATGGGTGGTCCTAGG B2.6 CGCCTGAATTTTCTGACTCTTCCCAT

C locus-specific primers

SGD008.CIVS3.R1NP ATCCCGGGAGATCTACAGGAGATG SGD011.CIVS3.R2NP AACAGCGCCCATGTGACCATCCT C2.1 CTGGGGAGGCGCCGCGTTGAGGATTCT C2.2 CGTCTCCGCAGTCCCGGTTCTAAAGTTCCCAGT C2.3 ATCCTCGTGCTCTCGGGA C2.4 TGTGGTCAGGCTGCTGAC C2.5 AAGGTTTGATTCCAGCTT C2.6 CCCCTTCCCCACCCCAGGTGTTCCTGTCCATTCTTCAGGA C2.7 CACATGGGCGCTGTTGGAGTGTCG

Class I loci-specific primers

SGD005.IIVS1.LNP GTGAGTGCGGGGTCGGGAGGGA 1.1 CACCCACCGGGACTCAGA 1.2 TGGCCCTGACCCAGACCTGGGC 1.3 GAGGGTCGGGCGGGTCTCAGC 1.4 CTCTCAGGCCTTGTTC 1.5 CAGAAGTCGCTGTTCC

DQA1 locus-specific primers

SGD001.DQA1.LNP TTCTGAGCCAGTCCTGAGA DQA3 E1a TTGCCCTGACCACCGTGATG DQA3 E1b CTTCCTGCTTGTCATCTTCA DQA3 E1c CCATGAATTTGATGGAGA DQA3 E1d ACCGCTGCTACCAATGGTA SGD003.DQA1.RNP CCAAGAGGTCCCCAGATC

DRA locus-specific primers

DRA E1 TCATCATAGCTGTGCTGATG DRA 5'E2 AGAACATGTGATCATCCAGGC DRA 3'E2 CCAACTATACTCCGATCACCAAT

DRB locus-specific primers

DRB E1 TGACAGTGACACTGATGGTGCTG DRB 5'E2 GGGGACACCCGACCACGTTTC DRB 3'E2 TGCAGACACAACTACGGGGTTG

DQB1 locus-specific primers

DQB E1 TGGCTGAGGGCAGAGACTCTCCC DQB 5'E2 TGCTACTTCACCAACGGGAC DQB 3'E2 GGTGTGCACACACAACTAC DQB 5'IVS1a AGGTATTTTACCCAGGGACCAAGAGAT DQB 5'IVS1b ATGTAAAATCAGCCCGACTGCCTCTTC DQB 3'IVS2 GCCTCGTGCCTTATGCGTTTGCCTCCT

DPB1 locus'specific primers

DPB E1 TGAGGTTAATAAACTGGAGAA DPB 5'IVS1 GAGAGTGGCGCCTCCGCTCAT DPB 3'IVS2 GAGTGAGGGCTTTGGGCCGG

What is claimed is:

1. A method for detection of at least one coding region allele of a multi-allelic genetic locus comprising:

a) amplifying genomic DNA with a primer pair that spans a non-coding region sequence, said primer pair defining a DNA sequence which is in genetic linkage with said genetic locus and contains a sufficient number of non-coding region sequence nucleotides to produce an amplified DNA sequence characteristic of said allele; and

b) analyzing the amplified DNA sequence to detect the allele.

2. The method of claim **1** wherein said amplified DNA sequence includes at least about 300 nucleotides corresponding to non-coding region sequences.

3. The method of claim **1** wherein said non-coding region sequence is adjacent to an exon encoding said allele.

4. The method of claim **1** wherein said amplified DNA sequence is characteristic of at least one nonadjacent allele.

5. The method of claim **1** wherein said amplified DNA sequence is characteristic of at least one adjacent allele and at least one nonadjacent allele.

6. The method of claim **5** wherein said amplified DNA sequence includes at least about 1,000 nucleotides corresponding to non-coding region sequences.

7. The method of claim **1** wherein said genetic locus has at least four alleles.

8. The method of claim **1** wherein said genetic locus has at least eight alleles.

9. A method for detection of at least one allele of a multi-allelic genetic locus comprising:

a) amplifying genomic DNA with a primer pair that spans a non-coding region sequence, said primer pair defining a DNA sequence which is in genetic linkage with said allele and contains a sufficient number of non-coding region sequence nucleotides to produce an amplified DNA sequence characteristic of said allele; and

b) analyzing said amplified DNA sequence to determine the presence of a genetic variation in said amplified sequence to detect the allele.

10. The method of claim **9** wherein said variation in said amplified DNA sequence is a variation in the length of the primer-defined amplified DNA sequence.

11. The method of claim **9** wherein said variation in said amplified DNA sequence is a change in the presence of at least one restriction site in the primer-defined amplified DNA sequence.

12. The method of claim **9** wherein said variation in said amplified DNA sequence is a change in the location of at least one restriction site in the primer-defined amplified DNA sequence.

13. The method of claim **9** wherein said variation in said amplified DNA sequence is a substitution of at least one nucleotide in the primer-defined amplified DNA sequence.

14. The method of claim **9** wherein said genetic locus is a major histocompatibility locus.

15. The method of claim **9** wherein said allele is associated with a monogenic disease.

16. The method of claim **15** wherein said monogenic disease is cystic fibrosis.

17. The method of claim **9** wherein at least about 70% of said primer-defined amplified DNA sequence corresponds to non-coding region sequences.

18. The method of claim **9** wherein said primer-defined amplified DNA sequence is from 300 to 500 nucleotides in length.

19. A method for producing RFLP patterns for an HLA locus of an individual comprising the steps of:

a) amplifying HLA DNA from said individual with a primer pair specific for said HLA locus under conditions suitable to produce an amplified DNA sequence, primer sites for said primers being located in intervening sequence I and in intervening sequence III when said HLA locus is a Class I locus and in intervening sequence I and in intervening sequence II when said locus is a Class II locus;

b) producing a digest by combining said amplified DNA sequence with at least one endonuclease that cleaves said amplified DNA sequence to yield a set of fragments having distinctive fragment lengths; and

c) producing RFLP patterns from said digest.

20. The method of claim **19** wherein said amplification comprises:

i) combining said primer pair with HLA DNA from said individual under hybridizing conditions for a period of time sufficient for each primer in said primer pair to produce an extension product which, when separated from its complement, can serve as a template for synthesis of the extension product of the other primer to produce a mixture;

ii) treating said mixture under denaturing conditions to separate the primers from their extension products;

iii) treating said mixture with said primer pair such that a primer extension product is synthesized using each of the templates produced in step ii as a template, resulting in amplification of the HLA DNA; and

iv) repeating steps ii and iii to produce an amplified DNA sequence.

21. The method of claim **19** wherein a second primer pair specific for said HLA locus is also used to amplify said HLA DNA.

22. The method of claim **19** wherein producing said RFLP fragment pattern comprises:

i). combining said amplified DNA sequence with at least one endonuclease that cleaves said amplified DNA sequence to yield a set of fragments having distinctive fragment lengths;

ii). separating said fragments based on the length of the fragments to produce separated fragments; and

iii). visualizing said separated fragments to produce RFLP fragment patterns.

23. The method of claim **22** wherein said fragments are separated using gel electrophoresis and visualized using a nucleotide-specific stain.

24. A method for determining whether DNA in a sample is from a particular individual comprising the steps of:

a) amplifying DNA from said individual and DNA from said sample with a primer pair specific for an HLA locus under suitable conditions to produce an amplified DNA sequence from said individual and from said sample, said primers being located in intervening sequences I and III for an HLA Class I locus and in intervening sequences I and II for a Class II locus;

b) combining said amplified DNA sequence from said individual and said amplified sample DNA from said sample with at least one endonuclease that cleaves said amplified DNA sequence into a plurality of cleaved sequences of sufficiently different lengths to distinguish between alleles of said HLA locus for a period of time sufficient for digestion of said amplified DNA to produce a digest; and

c) comparing restriction fragment length polymorphic patterns produced by said digest from said individual and from said sample to determine whether DNA in the sample is from the individual.

25. A method for determining whether an individual is the father of a child comprising the steps of:

a) amplifying DNA from said individual, DNA from said child and DNA from said child's mother with a pair of primers specific for an HLA locus under suitable conditions to produce amplified DNA sequences, said primers being located in intervening sequences I and III for an HLA Class I locus and in intervening sequences I and II for a Class II locus;

b) combining said amplified DNA sequence from said individual, said amplified DNA sequence from said child's mother, and said amplified sample DNA from said child with at least one endonuclease that cleaves said amplified DNA sequence into a plurality of cleaved sequences of sufficiently different lengths to distinguish between alleles of said HLA locus to produce a digest; and

c) comparing restriction fragment length polymorphic patterns produced by said digest from said individual, from said child's mother and from said child.

26. A DNA analysis method for determining coding region alleles of a multi-allelic genetic locus comprising identifying sequence polymorphisms characteristic of the alleles, wherein said sequence polymorphisms characteristic of the alleles are present in a non-coding region sequence, said non-coding region sequence being not more than about two kilobases in length.

27. The method of claim **26** wherein said sequence polymorphisms characteristic of the alleles are within five kilobases of a variable exon of the genetic locus.

28. The method of claim **26** wherein said sequence polymorphisms characteristic of the alleles are within two kilobases of a variable exon of the genetic locus.

29. The method of claim **26** wherein said sequence polymorphism characteristic of said coding region allele is present in an intervening sequence adjacent to a variable exon of the locus.

30. The method of claim **29** wherein the genetic locus is an HLA Class I locus and the intervening sequence is intervening sequence I, II or III.

31. The method of claim **29** wherein the genetic locus is an HLA Class II locus and the intervening sequence is intervening sequence I or II.

32. The method of claim **26** wherein said non-coding region sequence is not more than about one kilobase in length.

33. A method for producing RFLP fragments characteristic of alleles of an HLA locus of an individual comprising the steps of:

a) amplifying genomic HLA DNA from said individual with a primer pair specific for said HLA locus under conditions suitable to produce an amplified DNA sequence; and

b) producing a digest by combining said amplified DNA sequence with at least one endonuclease that cleaves said amplified DNA sequence to yield a set of fragments having distinctive fragment lengths.

34. The method of claim **33** additionally comprising the step of producing RFLP patterns from said digest.

35. The method of claim **33** wherein said primers define a DNA sequence that contains all exons that encode allelic variability associated with said HLA locus.

36. A method for producing RFLP fragments for an HLA locus of an individual comprising the steps of:

a) amplifying genomic HLA DNA from said individual with a primer pair specific for said HLA locus under conditions suitable to produce an amplified DNA sequence, said primers defining a DNA sequence that contains all exons that encode allelic variability associated with said HLA locus; and

b) producing a digest by combining said amplified DNA sequence with at least one endonuclease that cleaves said amplified DNA sequence to yield a set of fragments having distinctive fragment lengths.

* * * * *

US005612179C1

(12) EX PARTE REEXAMINATION CERTIFICATE (7478th)

United States Patent

Simons

(10) Number: US 5,612,179 C1

(45) Certificate Issued: May 4, 2010

(54) **INTRON SEQUENCE ANALYSIS METHOD FOR DETECTION OF ADJACENT AND REMOTE LOCUS ALLELES AS HAPLOTYPES**

(75) Inventor: **Malcolm J. Simons**, Fryerstown (NZ)

(73) Assignee: **Genetic Technologies Limited**, Fitzroy, Victoria (AU)

Reexamination Request:
No. 90/010,318, Oct. 15, 2008

Reexamination Certificate for:

Patent No.:	**5,612,179**
Issued:	**Mar. 18, 1997**
Appl. No.:	**07/949,652**
Filed:	**Sep. 23, 1992**

Related U.S. Application Data

(63) Continuation of application No. 07/551,239, filed on Jul. 11, 1990, now Pat. No. 5,192,659, which is a continuation-in-part of application No. 07/465,863, filed on Jan. 16, 1990, now abandoned, which is a continuation-in-part of application No. 07/405,499, filed on Sep. 11, 1989, now abandoned, which is a continuation-in-part of application No. 07/398,217, filed on Aug. 25, 1989, now abandoned.

(51) **Int. Cl.**

C12Q 1/68	(2006.01)
C12P 19/34	(2006.01)
C07H 21/04	(2006.01)
C12N 15/00	(2006.01)

(52) **U.S. Cl.** **435/6**; 435/91.1; 435/91.2; 536/23.1; 536/24.3; 536/24.31; 536/24.33

(58) **Field of Classification Search** ..... None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,302,204	A	11/1981	Wahl et al.
4,458,066	A	7/1984	Caruthers et al.
4,582,788	A	4/1986	Erlich
4,683,194	A	7/1987	Saiki et al.
4,683,195	A	7/1987	Mullis et al.
4,683,202	A	7/1987	Mullis
4,772,549	A	9/1988	Frossard
4,835,098	A	5/1989	Orr et al.
4,994,370	A	2/1991	Silver et al.
5,075,217	A	12/1991	Weber
5,192,659	A	3/1993	Simons
5,200,313	A	4/1993	Carrico
5,310,893	A	5/1994	Erlich et al.
5,364,759	A	11/1994	Caskey et al.
5,612,179	A	3/1997	Simons
5,789,568	A	8/1998	Simons
5,851,762	A	12/1998	Simons
6,194,147	B1	2/2001	Baxter-Lowe et al.
6,503,707	B1	1/2003	Baxter-Lowe
2002/0197613	A1	12/2002	Canck et al.
2003/0119003	A1	6/2003	Simons
2004/0197775	A1	10/2004	Simons et al.

FOREIGN PATENT DOCUMENTS

AU	16727/88	12/1988
AU	36908/89	12/1989
AU	45565/89	5/1990
DE	69029018	5/1997
EP	0237362	9/1987
EP	0256630	2/1988
EP	0269260	6/1988
EP	0364255	4/1990
EP	0370719	5/1990
EP	0414469	2/1991
EP	0443748	8/1991
WO	WO 89/12697	12/1969
WO	WO 89/04168	5/1989
WO	WO 89/04875	6/1989
WO	WO 89/11547	11/1989
WO	WO 90/02818	3/1990
WO	WO 90/12891	11/1990

OTHER PUBLICATIONS

Funke et al., "Detection of a New MSP I Restriction Fragment Length Polymorphism in the Apolipoprotein A–1 Gene," J. Clin Chem. Clin. Biochem., 25(3): 131–134, 1987.*

Geisel et al., "A New Aga L1 Restriction Fragment Length Polymorphism in the Low Density Lipoprotein Receptor Gene," J. Clin. Chem. Clin. Biochem., 26(7): 429–434, 1988.*

Koller et al., "Cloning and Complete Sequence of an HLA–A2 Gene: Analysis of Two HLA–A Alleles at the Nucleotide Level," J. Immunology, 134(4): 2727–2733, Apr., 1985.*

Koller et al., "Comparison of Multiple HLA–A Alleles at the DNA Level by Using Southern Blotting and HLA–A–Specific Probes," J. of Immunology, 135(6): 4229–4234, Dec. 1985.*

N'Guyen et al., "The HLA–AW24 Gene: Sequence, Surrounding and Comparison with HLA–A2 and HLA–A3 Genes," Immunogenetics, 21, 479, 1985.*

(Continued)

Primary Examiner—Gary L Kunz

(57) **ABSTRACT**

The present invention provides a method for detection of at least one allele of a genetic locus and can be used to provide direct determination of the haplotype. The method comprises amplifying genomic DNA with a primer pair that spans an intron sequence and defines a DNA sequence in genetic linkage with an allele to be detected. The primer-defined DNA sequence contains a sufficient number of intron sequence nucleotides to characterize the allele. Genomic DNA is amplified to produce an amplified DNA sequence characteristic of the allele. The amplified DNA sequence is analyzed to detect the presence of a genetic variation in the amplified DNA sequence such as a change in the length of the sequence, gain or loss of a restriction site or substitution of a nucleotide. The variation is characteristic of the allele to be detected and can be used to detect remote alleles. Kits comprising one or more of the reagents used in the method are also described.

OTHER PUBLICATIONS

Stetler et al., "Polymorphic restriction endonucleases sites linked to the HLA–DRa gene: Localization and use as genetic markers of insulin–dependent diabetes," Proc. Natl. Acad. Sci. 82: 8100–8104, Dec., 1985.*

Koller et al., "Cloning and Complete Sequence of an HLA–A2 Gene: Analysis of Two HLA–A A1 at the Nucleotide Level," J. Immunol. 133: 2727. 1985.*

N'Guyen et al., "The HLA–AW24 Gene: Sequence, Surroundings and Comparison with the HLA–A2 and HLA–A3 Genes," Immunogenetics 21: 479, 1985.*

Stetler et al., "Polymorphic restriction endonuclease sites linked to the HLA–DRalpha gene: Localization and use as genetic markers of insulin–dependent diabetes," Proc. Natl. Acad. Sci. 82: 8100–8104, 1985.*

Götze, et al., "The Major Histocompatibility Complex", Fundamentals of Immunology, Springer Verlag, 1981 (ed. Bier, da Silva, Götze, Mota).

"Conserved Sequences at Exon–Intron Boundaries in Pre–mRNA," Molecular Biology of the Gene, 4th Edition, California, 1987, Chapter 20, pp. 639–640.

Molecular Biology of The Cell, 2nd Edition, New York, 1989, Chapter 9: "RNA Synthesis and RNA Processing" and Chapter 10: Control of Gene Expression, pp. 523–612.

Royer–Pokora, et al. "Cloning the Gene for the Inherited Disorder Chronic Granulomatous Disease on the Basis of Its Chromosomal Location" Cold Sp. Harbor Symp. LI:177–183 (1986).

Lebo, et al. "Flow–sorting Analysis of Normal and Abnormal Human Genomes" Cold Sp. Harbor Symp. LI:169–176 (1986).

U.S. Appl. No. 07/398,217, filed Aug. 25, 1989, Simons.

U.S. Appl. No. 07/405,499, filed Sep. 11, 1989, Simons.

U.S. Appl. No. 07/465,863, filed Jan. 16, 1990, Simons.

U.S. Appl. No. 07/550,939, filed Jul. 11, 1990, Simons.

U.S. Appl. No. 07/907,721, filed Nov. 8, 1991, Simons.

U.S. Appl. No. 07/971,856, filed Mar. 9, 1993, Simons.

U.S. Appl. No. 09/070,497, filed Apr. 30, 1998, Simons.

A.S.E.A.T.T.A. "Donor Reactive Alloantibodies Auto–Antibodies Revisited Aseatta Quality Control Transplantation in Thailand", vol. 20, No. 2, 1995, Newsletter, 15 pages.

Adams, et al Sci. 252:1651–1656 (1991).

Amselem, et al., "Determination of the Spectrum of β–Thalassemia Genes in Spain by Use of Dot–Blot Analysis of Amplified β–Globin DNA," Am. J. Hum. Genet, 43:095–100, 1988.

Arnheim et al., "Use of Pooled DNA Samples To Detect Linkage Disequilibrium of Polymorphic Restriction Fragments and Human Disease: Studies of the HLA Class II Loci," Proc. Natl. Acad. Sci. USA 82:6970–6974 (Oct. 1985).

Baird, et al. A Nucleotide Change At A Spice Junction In The Human B–Globin Gene is Associated With B–Thalassemia. Jul. 1981. pp. 4218–4221. vol. 78 No. 7. Proc. Natl. Acad. Sci. USA.

Boehnke et al., "Fine–Structure Genetic Mapping of Human Chromosomes Using the Polymerase Chain Reaction on Single Sperm: Experimental Design Considerations," The American Journal of Human Genetics, vol. 45, No. 1, Jul. 1989, pp. 21–32.

Boerwinkle et al, PNAS, 86:212–216 (1989).

Brinster, et al. Proc. Natl. Acad. Sci. (1988) 85: 836–840.

Brooks, et al., "Association of the Widespread A149P Hereditary Fructose Intolerance Mutation with Newly Identified Sequence Polymorphisms in the Aldolase B Gene," Am. J. of Hum. Genet., 52:835–840 (1993).

Bugawan et al., "The Use of Non–Radioactive Oligonucleotide 'Probes to Analyze Enzymatically Amplified DNA for Prenatal Diagnosis and Forensic HLA Typing", Biotechnology 6:943–947 (1988).

Bugawan et al., J. Immunol. vol. 141, No. 12, pp. 4024–4030, 1988.

C.K. Hurley, et al. "Large–Scale DNA–Based Typing of HLA–A and HLA–B At Low Resolution Is Highly Accurate Specific and Reliable", Tissue Antigens 55: 352–358, 2000.

Cavalli–Sforza, L. Am. J. Hum. Genet. 46:649–651 (1990).

Cereb, Nezih et al., "Locus–Specific Conservation of the HLA Class Introns by Intra–Locus Homogenization," (1997) Immunogenetics, 47: 30–36.

Certified Translation of Inoko (1989) "DNA Typing of HLA Antigen," Nihon Rinsho, 47, 550–576.

Chien et al, J. Bacteriol., 127:1550–1557 (1976).

Clark, "Interference of Haplotypes from PCR–amplified Samples of Diploid Populations", Mol. Biol. Evol., vol. 7, No. 2, 1990, pp. 111–112.

Coleman. et al. "Polymorphisms in The Apolipoprotein AIOCIII Gene Complex." 1986. pp. 213–228. Mol. Biol. Med. (1986) 3, 1986 Academic Press Inc. (London) Ltd.

Coleman. et al. Chemical Abstracts. Sep. 15, 1986. pp. 175–176. vol. 105 No. 11. Chemical Abstracts Service–The Americal Chemical Society.

Collins et al., "Variations on a Theme: Cataloging Human DNA Sequence Variation" Science Nov. 28, 1997: vol. 278. No. 5343, pp. 1580–1581.

Caruthers, Science 230:281–285 (1985).

Degli–Esposti, et al., "Updated Characterization Of The Fourth Ancestral Haplotypes Using Asia–Ocenia Histocompatibility Workshop Panel", Human Immunology 44, 12–18 (1995).

Den Dunnen et al, Am. J. Hum. Genet. 45(6):835–847 (1989).

Deng, "A sensitive non–radioactive PCR–RFLP analysis for detecting point mutations at 12th condon oncogene c–Ha–ras in DNAs of gastic cancer," Nucleic Acids Research, vol. 16. No. 13, 1988, p. 6231.

Di Marzo, et al., "The Spectrum of β–Thalassaemia Mutations in Sicily," British Journal of Haematology, 69, 393–397, 1988.

Dicker, et al., "Sequence Analysis of a Human Gene Responsible for Drug Resistance: A Rapid Method for Manual and Automated Direct Sequencing of Products Generated by the Polymerase Chain Reaction," BioTechniques, vol. 7, No. 8, 1989, pp. 830–837.

Diethard Tauiz, "Hypervariability of Simple Sequences As A General Source For Polymorphic DNA Markers", vol. 17, No. 16, 1989.

DiLella, et al., "Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria," Nature, vol. 322, No. 28, Aug. 1986, pp. 799–803.

Duyk, et al. Proc. Natl. Acad. Sci. (1990) 87:8995–8999.

Eisses, et al., "Analysis of the Gene Encoding the Multifunctional Alcohol Dehydrogenase Allozyme ADH–71k of Drosophila Melanogaster," Mol. Biol. Evol., 7:459–469 (1990).

Erlich et al., Biotechnology 4:975–981 (1986).

Feder et al. "A Novel MHC class I–likke gene is mutated in patients with hereditary haemochromatosis", Nature Genetics, vol. 13, Aug. 199, pp. 399–408.

Funke et al., Biosis No. 84025050 Abstract.

Funke. et al., "Pst I RFLP close to the LDL receptor gene—Synthetic 45 Base Oligodeoxynuclectide Representing A Sequence From the 18th Exon of The LDL Receptor Gene." p. 7820. vol. 14 No. 19.IRL Press Umited, Oxford. England, Nov. 19, 1986.

Gavrilidis, A., et al., "HLA DRB1 Locus Sequence Typing", Genetype, Melbourne, Victoria, date unknown, pp. 1–2.

Geisel et al., Abstract Biosis No. 87082065.

Gitschier et al. "Genetic mapping and diagnosis of haemophilia A achieved through a Bcll polymorphism in the factor VIII gene", Nature. 1985 Apr. 25–May 1;314(6013):pp. 738–740.

Goldman et al, Electrophoresis, 3:24–26 (1982).

Graham et al., "Application of Molecular Genetics to Prenatal Diagnosis and Carrier Detection in the Hemophilias: Some Limitations," Blood, vol. 66, No. 4, Oct. 1985, pp. 759–764.

Green et al., "Chromosomal Region of the Cystic Fibrosis Gene in Yeast Artificial Chromosomes: A Model for Hwnan Genome Mapping" Science, vol. 250, pp. 94–98.

Grimm et al, Am. J. Hum. Genet. 45(3):368–372 (1989).

Guardiola. John et al., "Molecular Genotyping of the HLA–DQalpha Gene Region." (1988) Immunogenetics,27: 12–18.

Gyllensten et al., Proc. Natl. Acad. Sci. USA, vol. 85, pp. 7652–7656, 1988.

H. J. Noreen et al., "Validation Of DNA–Based HLA–A and HLA–B Testing Of Volunteers For A Bone Marrow Registry Through Parallel Testing With Serology", Tissue Antigens 57: 221–229, 2001.

Harano et al., "Hemoglobin Abnormalities In A Black Family With HB S, Hereditary Persistence Of HB F, and A Y Chain Variant; A Reevaluation Through Gene Mapping", Hemoglobin, vol. 8(6), pp. 549–568, (1984).

Horn et al., "Characterization and Rapid Diagnostic Analysis of DNA Polymorphisms Closely Linked to the Cystic Fibrosis Locus," Clin. Chem. 36: 1614–1619 (1990).

Hyldig–Nielsen et al., Proc. Natl. Acad. Sci, USA 84:1644–1648 (1987).

Inoko (1989) "DNA Typing of HLA Antigen," Nihon Rinsho, 47, 550–576.

Iwahana, et al., "A New RFLP in Intron 1 of the Human c–Ha–ras1 Gene and its Close Relationship with the Variable Tandem Repeats in the Region 3' to the Gene," Oncogene, 5:1049–53, 1990.

Jeffreys et al., "Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells", Nucleic Acid Research 16:10953–10971 (1988).

Jeffreys et al., Nature, 314 (1985), pp. 67–73.

Kan et al, N. Engl. J. Med. 297: 1081–1084 (1977).

Kazazian, Jr., Haig H. et al., "Molecular Basis and Prenatal Diagnosis of P–Thalassemia," (1988) The Journal of The American Society of Hematology, vol. 72, No. 4, pp. 1107–1116.

Kerem et al., "Identification of the Cystic Fibrosis Gene: Genetic Analysis", Science, vol. 245, 1989, pp. 1073–1080.

Kogan et al., "An Improved Method for Prenatal Diagnosis of Genetic Diseases by Analysis of Amplifed DNA Sequences—Application to Hemophilia A," The New England Journal of Medicine, vol. 317, No. 16, Oct. 15, 1987, pp. 985–990.

Kotsch, K. et al., "Sequencing of HLA Class I Genes Based on the Conserved Diversity of the Noncoding Regions: Sequencing–Based Typing ofthe HLA–A Gene," (1997) Tissue Antigens, 50: 178–191.

Lander et al., "Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children", Science 236: 1567–1570, Jun. 1987.

Lawrence, et al., "Neonatal Lupis In Twins", Southern Medical Journal, vol. 82, No. 5, 1989, pp. 657–660.

Liebhaber, et al., "Compensatory Increase in α 1–Globin Gene Expression in Individuals Heterozygous for the α–Thalassemia–2 Deletion," J. Clin Invest, Sep. 1985, 76(3): 1057–1064.

Limm et al., "HLA–DQAI Allele and Suballele Typing Using Noncoding Sequence Polymorphisms Application to 4AOHW Cell Panel Typing", Human Immunology 38, 57–68 (1993).

Little et al., "Nature", 285: 144–147, 1980.

Maeda et al., Tissue Antigens, (1989), 34:290–298.

Mariapia A. Degli–Esposti, et al., Characterization of 4AOHW Cell Line Panel Including New Data For The 10IHW Panel, Hum. Immunol. 1993,38: 3–16.

Marshall, Electrophoresis, 4:269–272 (1983).

Marx, "Disecting the Complex Diseases," Science, vol. 247, Mar. 30, 1990, pp. 1540–1542.

McGinnis et al., Ancient, Highly Polymorphic Human Major Histocompatibility Complex DQA1 Intron Sequences, American Journal of Medical Genetics, 52:438–444 (1994) © 1994 Wiley–Liss, Inc.

Messer, et al., "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An Ncol Polymorphism in the First Intron of the Human TNF–âGene Corrolates with Variant Amino Acid in Position 26 and a Reduced Level of TNA âProduction," J. Exp. Med., 173: 209–219 (1991).

Myerowitz, "Splice Junction Mutation in Some Ashkenazi Jews with Tay–Sachs Disease Evidence Against a Single Defect Within this Ethnic Group," Proc. Natl. Acad. Sci. USA, Jun.1988 85(11): 3955–59.

Nakamura, et al. Science (1987) 235: 1616–1620.

Nakatsuji, et al., "Fetal Hemoglobin Variants Identified in Adults Through Restriction Endonuclease Gene Mapping Methodology", Blood, vol. 66, No. 4, 1985, pp. 803–807.

Nathans et al., Annual Review of Biochemistry 44:273–293 (1975).

Naughton et al., DPBI Locus PCR–RFLP Typing Of The Fourth Asia–Oceania Histocompatibility Workshop Cell Panel Reveals A Novel OPBI Allele, European Journal Of Immunogentics (1994), 21, 351–364.

Nickerson et al. "Identification of Clusters of Biallelic Polymorphic Sequence–Taggeg sites (pSTSs) That Generate Highly Informative and Automatable Markers for Genetic Linkage Mapping", Genomics 12, 377–387, 1992.

Ochman et al., "Genetic Applications of an Inverse Polymerase Chain Reaction", Genetics 120: 621–623 (1988).

Olson et al, Science 245: 1434–1435 (1989).

Orkin, et al, "Nonsense And Frameshift Mutations in β–Thalassemia Detected in Cloned β–Globin Genes." Oct. 10. 1981. pp. 9782–9784. vol. 256. No. 19. The Journal Of Biological Chemistry USA.

Paul, et al, "DNA Polymorphic Patterns and Haplotype Arrangements of the apo A–I, apo C–ID, apc A–IV Gene Cluster in Different Ethnic Groups," Hum Genet (1987) 75: 264–268.

R. V. Lebo, et al Am. J. Hum. Genet 47:583–590 (1990).

Ricketts, et al, "Defective Splicing of Thyroglobulin Gene Transcripts in the Congenital Goitre of the Afrikander Cattle," The EMBO Journal, vol. 4, No.3, pp. 731–737, 1985.

Rieβ, et al., "Hypervariability of intronic simple (gt).sub.n (ga).sub.m repeats in HLA–DRB genes," Immunogenetics, vol. 32, No. 2, Aug. 1990, pp. 110–116.

Rigby et al, J. Mol. Biol., 113: 237 (1977).

Riordan et al., "Identification of the Cystic Fibrosis Gene: Cloning and Characterization of Complementary DNA,", Science, vol. 245, 1989, pp. 1066–1073.

Roger L. Dawkins, "Conclusions And Proposals For Future Workshops", Human Immunology, vol. 38, pp. 83–85, (1993).

Rommens et al., "Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping,", Science, vol. 245, 1989, pp. 1059–1065.

Ruano et al, "Direct haplotyping of chromosomal segments from multiple heterozygotes via allele–Specific PCR amplification," Nucl. Acids. Res. 17:8392 (1989).

Ruano. et al., "Haplotype Of Multiple Polymorphisms Resolved By Enzymatic Amplification Of Single DNA Molecules." Aug. 1990. pp. 6296–300. vol. 87. Proc. Natl. Acad. Sci USA.

Saiki et al., "Enzymatic Amplification of .beta–Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia,", Science, vol. 230, 1985, pp. 1350–1354.

Saiki, Gyllensten and Erlich, The Polymerase Chain Reaction in Genome Analysis: A Practical Approach, ed. Davies pp. 141–152, (1988) I.R.L. Press, Oxford.

Sallee et al., Am, J. Human Genet., 45(4): A216, 1989.

Sargent et al EMBO 8:2305–2312 (1989).

Scharf et al., Science, Sep. 5, 1986, 233(4768), pp. 1076–1078.

Scheffer et al, Am. J. Hum. Genet. 45(2):252–260 (1989).

Schwartz et al., "Ordered Restriction Maps of Saccharomyces carevisiae Chromosomes Constructed by Optical Mapping", Science, vol. 262, Oct. 1, 1993, pp. 110–114.

Simons, et al., "HLA Haplotyping Using Non–Coding Polymorphisms", 4th Chromosome 6 Workshop, (1999), 4 pages.

Simons, et al., "Strategy For Definition Of Drug Haplotypes In The 4AOHW Cell Panel Using Noncoding Sequence Polymorphisms", Human Immunology, vol. 38, pp. 69–74 (1993).

Singer et al., "Genes & Genomes", 1991, pp. 698–699, 873, 929.

Skolnick et al (1989) "Simultaneous analysis of multiple polymorphic loci, using amplified sequence polymorphisms (ASPo)." Genomics 2:273–279.

Smeets et al., Use Of Variable Simple Sequence Motifs As Genetic Markers: Application To Study Of Myotonic Dystrophy, Hum Genet (1989) 83: 245–251.

Stephens et al., "Theoretical Underpinning of the Single–Molecule–Dilution (SMD) Method of Direct Haplotype Resolution," Am. J. Hum. Genet, vol. 46, 1990, pp. 1149–1155.

Taylor et al, Nature 251:392–393 (1974).

Tegelstrom, Electrophoresis, 7:226–229 (1986).

The Jornal of the Japan Medical Association, vol. 97 (1987), pp. 637–640.

Traver, et al. Proc. Natl. Acad. Sci. (1989) 86: 5898–5902.

Varney, A Gavrilidis, B.D. Tait, "Polymorphism In The Regulatory Regions Of HLA–DPB1" Tissue Laboratory, Department Of Pathology, The Royal Melbourne Hospital, Grattan Street, Parkville 3050, Australia,. date unknown, 1 page.

Wahl, et al, PNAS 6:3683–3687 (1979).

Waikan et al. P.N.A.S, U.S.A., vol. 75, No. 11, pp. 5631–5635, Nov. 1978.

Wallace, et al Sci. 249:181–186 (1990).

Weber et al. (1989) "Abundant class of human DNA polymorphisms which can be typed using polymerase Chain Reaction," Am. J. Hum. Gen. 44: 388–396.

White, et al, "The Polymerase Chain Reaction," TIG, Jun. 1989, vol. 5, No. 6. pp. 185–189.

Wong. Corrine et al., "Characterization of β–thalassaemia Mutations Using Direct Genomic Sequencing of Amplified Single Copy DNA," (1987) Nature, vol. 330, pp. 384–386.

Zeng, et al., "Hereditary Persistence Of Fetal Hemoglobin Or (ōβ)°—Thalassemia: Three Types Observed In South–Chinese Families", Blood, vol. 66, No. 6, 1985, pp. 1430–1435.

Geisel, J. Clin. Chem. Clin. Biochem., 26(7), 1988, 429–434.

Aseatta, 1996 Annual Scientific Meeting, M.J. Simons, Genetype P/L. Fitzroy, Victoria, 3065.

David Botstein. et al., Construction Of A Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms, .1980. pp. 314–331. Am J. Hum Genet 32. 1980 American Soicety of Human Genetics.

Elsevier, "Human Immunology", Fourth Asia–Oceania Histocompatibiutyworkshop, Apr. 22–29, 1993. Perth, Western Australia, 1993HUIMDQ 38(1), 1993, pp.1 , 2, 57–85.

Simons, et al., "Donor Reactive Alloantibodies Auto–Antibodies Revisted Aseatta Quality Control Transplantation In Thailand", vol. 20, No.2 (1995).

Basic & Clinical Immunology, 3rd Ed (1980) Lange Medical Publications, pp. 181–190.

Abraham, et al., "Sequence differences between HLA–B and TNF distinguish different MHC ancestral haplotypes," Tissue Antigens, vol. 39, No. 3, Mar. 1992, pp. 117–121.

Akerman, et al., "Identification of Deletion and Triple α–Globin Gene Haplotypes in the Montreal β–Thalassemia Screening Program: Implications for Genetic Medicine," Am. J. Med. Genet., 36:76–84 (1990).

Allen et al., Biotechniques, vol. 7, No. 7, pp. 736–744 (1989).

Blasczyk et al., "Sequence analysis of the 2nd intron revealed common sequence motifs providing the means for a unique sequencing based typing protocol of the HLA–A locus." Tissue Antigens. 1996 Feb; 47(2):102–10.

Brown et al, Meth. Enzymol., 68:109–151 (1979).

Burnside, et al, Endocrinology (1991) 128: 3183–3191.

Carmi et al., "Use of a DNA pooling strategy to identify a human obesity syndrome locus on chromosome 15", Human Molec. Genet. 4: 9–13, 1995.

Compton Nature 350:91 [1991].

Cremer, et al. Hum. Genet. (1988) 80: 235–246.

DiLella, et al., "Screening for Phenylketonuria Mutations by DNA Amplification with the Polymerase Chain Reaction," The Lancet, Mar. 5, 1988, pp. 497–499.

Erlich et al., "Reliability of the HLA–DQ.alpha. PCR–Based Olignucleotide Typing System," Journal of Forensic Sciences, Sep. 1990, pp. 1017–1018.

Fries, et al., "The Bovine Genome Contains Polymorphic Microsatellites," Genomics, Oct. 1990; 8(2):403–6.

Goto et al., "Genetic linkage of Werner's syndrome to five markers on chromosome 8", Nature 355: 735–738, Feb. 20, 1992.

Honig, et al., "Human Hemoglobin Genetics," 1986, Springer–Verlag, pp. 227–249.

Mardis et al., "Automated Methods for Single–Stranded DNA Isolation and Dideoxynucleotide DNA Sequencing Reactions on a Robotic Workstation," BioTechniques, vol. 7, No. 8, 1989, pp. 840–850.

Monaco, et al, Nature 323:656–650 (1986).

Nrang, et al, Meth. Enzymol., 68:90–98 (1979).

O'Hara et al. (1988) "The human factor VII gene is polymorphic due to variation in repeat copy number in a minisatellite," Gene 66: 147–158.

Padgett et al., "Splicing of Messenger RNA Precursors," Ann. Rev. Biochem., 55:1119–1150 (1986).

Wyngaarden, Human Gene Therapy 1:73–92 (1990).

S. H. Orkin, et al Ann. Rev. Genet 18:131–171 (1984).

Saiki et al., Nature, vol. 324, pp. 163–166, 1986.

Schreuder et al., Human Immunology, Nov. 1999, vol. 60, pp. 1157–1181.

Schreuder et al., Tissue Antigens, Aug. 2001, vol. 58, pp. 109–140.

Sheng–Jing Lu, et al., "Linkage Of A Nasopharyngeal Carcinoma Susceptibility Locus To The HLA Regions", Letters To Nature, Nature 346,470–471 (1990).

Southern, J. Mol. Biol., 98:503–517 (1975).

Zhang et al. (1987) Structure of the mouse Adh–1 gene etc. Gene 57: 27–36.

Jeffreys "DNA Sequence Variants in the Gy–, Ay–, Ōy–and β–Globin Genes of Man" Cell, vol. 18, pp. 1–10, Sep. 1979.

Bröcker–Vriends et al., "Carrier Detection of Haemophilia B by Using Intragenic Restriction–Fragment Length Polymorphism", Thrombosis and Haemostasis, 54(2), pp. 506–509, 1985.

Connor et al., "Application of three intragenic DNA polymorphisms for carrier detection in haemophilia B", Journal of Medical Genetics 1986, vol. 23, pp. 300–309.

Rees et al., "Haplotypes identified by DNA polymorphisms at the apolipoprotein A–1 and ClII loci and hypertriglyceridaemia", Hum. Genet (1986), vol. 72, pp. 168–171.

Moodie et al., "Carrier detection in 50 haemophilia A kindred by means of three intragenic and two extragenic restriction fragment length polymorphisms", British Journal of Haematology, 1988, vol. 70, pp. 77–84.

"Linkage disequllibrium", copy from Wikipedia, printed Jan. 9, 2008, pp. 1–4.

Strobeck, Genetics 103, pp. 545–555 (1983).

Casula et al., Blood, vol. 75, No. 3, p. 662–670 (1990).

Beaucage et al, Tetrahedron letters, 22:1859–1862 (1981).

Maniatis et al. "Molecular Cloning A Laboratory Manual", 1982, p. 437.

Simons and Erlich, pp. 952–959 In: Immunology of HLA, vol. 1: Springer–Verlag, New York (1989).

Simons et al., pp. 959–1023 In: Immunology of HLA, vol. 1: Springer–Verlag, New York (1989).

New England BioLabs 1988–1989 Catalog, p. 62, 1988.

Search Report for nucleotide residue sequence "CTCTCAGGCCTTGTT", Dec. 1996.

Genetic Tecnologies Limited vs. *Applera Corporation,* "Reply of Genetic Technologies Limited To Counterclaims of Applera Corporation" Sep. 25, 2003.

Genetic Technologies Limited vs. *Covance, Inc., And Variagenics Inc. Now Known As Nuvelo Inc., And Nuvelo, Inc.,* "Answer And Counterclaims By Covance, Inc., and Nuvelo, Inc., To Plaintiffs Complaint For Patient Infringement Demand For Jury Trial" Aug. 26, 2003.

Genetic Technologies Ltd., v. *Applera Corporation,* "Claim Construction Order", dated Sep. 15, 2004, pp. 1–13.

Genetic Technologies Ltd., v. *Applera Corporation,* "Answer and Counterclaims of Defendant Applera Corporation", dated Sep. 5, 2003, pp. 1–16.

Genetic Technologies Limited vs. *Applera Corporation,* "Complaint For Patent Infringement" Mar. 26, 2003.

"Nullification Action of Bioscientia Institute for Medical Diagnostics" against EP–B1 0 414 469 (German patent 690 29 018), Jun. 21, 2007, 860 pages.

Translation of Opposition dated Jul. 15, 2007 against the nullity complaint, 136 pages.

Translation of Reply to the brief of the patent proprietor dated Dec. 13, 2007 supplying grounds for its objection, pp. 1–28.

Schlieben, et al., "Genotyping of Bovine Kappa–Casein (Kappa–CAN, Kappa–CNB, Kappa–CNC, Kappa–CNE) Following DNA Sequence Amplification and Direct Sequencing of Kappa–CNE PCR Product," Anim. Genet. 1991; 22(4):333–42.

Seilhamer et al., DNA, vol. 3, No. 4, 309–317 (1984).

* cited by examiner

EX PARTE REEXAMINATION CERTIFICATE ISSUED UNDER 35 U.S.C. 307

NO AMENDMENTS HAVE BEEN MADE TO THE PATENT

AS A RESULT OF REEXAMINATION, IT HAS BEEN DETERMINED THAT:

The patentability of claims **26–32** is confirmed.

Claims **1–25** and **33–36** were not reexamined.

* * * * *